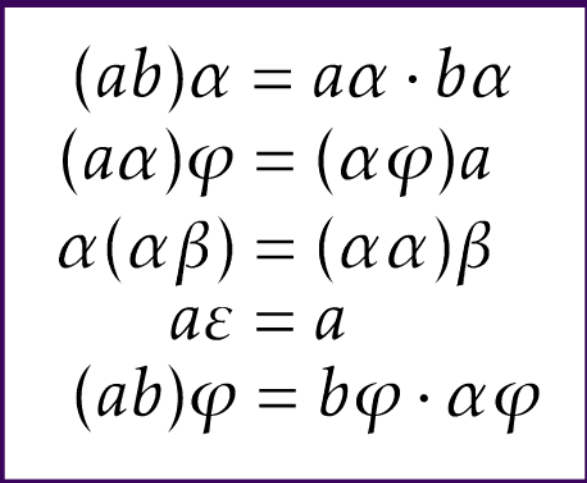
A. G. KUROSH

LECTURES ON GENERAL ALGEBRA



LECTURES ON GENERAL ALGEBRA

BY

A. G. KUROSH

AUTHORIZED TRANSLATION
BY
K. A. HIRSCH

CHELSEA PUBLISHING COMPANY
NEW YORK, N. Y.

THE PRESENT WORK IS AN ENGLISH TRANSLATION, BY PROFESSOR K. A. HIRSCH, OF THE RUSSIAN-LANGUAGE BOOK LEKCII PO OBSHCHEI ALGEBRE BY PROFESSOR A. G. KUROSH (А. Г. Курош)

COPYRIGHT ©, 1963 BY CHELSEA PUBLISHING COMPANY

COPYRIGHT, 1963, BY CHELSEA PUBLISHING COMPANY

©, 1963, CHELSEA PUBLISHING COMPANY

LIBRARY OF CONGRESS CATALOGUE CARD NO. 63-15996

TYPESETTER: A. ZELNICK

PRINTED IN THE UNITED STATES OF AMERICA

AUTHOR'S PREFACE TO ENGLISH EDITION

GENERAL ALGEBRA is represented in the present-day mathematical literature by a large number of books that differ both in the choice of material and in the character of the exposition. I believe that a book of the type of my *Lectures on General Algebra* has an appropriate place among them.

The publication of an English translation of this book is my second contact, as author, with the English-speaking mathematical public: the first was the publication of the English translation of my *Theory of Groups*.

These contacts give me much pleasure, and I am very grateful to Professor K. A. Hirsch, who has assumed the heavy burden of translating both these books, and also to the Chelsea Publishing Company, which has brought them to light.

Moscow, January 1963 A. G. KUROSH

INTRODUCTION

IN THE COURSE of the twenties and thirties of the present century, a wide circle of mathematicians became aware of a radical reorganization that had taken place in algebra, one of the oldest branches of mathematics. This reorganization — specifically, the transformation of algebra into a set-theoretical axiomatic science having as its primary object of study the algebraic operations performed on elements of an arbitrary nature—had, of course, been prepared for by the whole preceding development of algebra. It began at the end of the nineteenth century and continued with increasing momentum during the first decades of the twentieth. But it was only the publication, in 1930 and 1931, of the two volumes of van der Waerden's *Moderne Algebra* that made the ideas, results, and methods of this 'new' algebra accessible to all mathematicians, including non-algebraists.

How great, and sometimes decisive, the impact of this modern algebra was on the development of many domains of mathematics, among which we mention, in the first instance, topology and functional analysis, is common knowledge. At the same time, the last three decades have witnessed an intensive, even stormy development of algebra itself, as well as the discovery of several new links with neighboring disciplines; as a result, modern algebra—or general algebra, as we prefer to call it—now presents an altogether different appearance from that of thirty years ago.

During these decades, the older branches of general algebra—the theory of fields and of associative and associative-commutative rings, to which van der Waerden's book was mainly devoted—have undergone far-reaching changes. Even more crucial was the reorientation in the theory of groups, the oldest of all the branches of general algebra. At the same time, the theory of rings became more and more a theory of non-associative rings, incorporating as a constituent part the theory of Lie rings and Lie algebras. Topological algebra sprang up and soon occupied a very prominent position, and a parallel development took place in the theory of ordered algebraic structures. The theory of lattices made its appearance and developed rapidly; and the last few years have seen the rise of the parallel theory of categories, which undoubtedly has a most important future. Within the

framework of the classical parts of general algebra, independent trends arose: homological algebra, which has already led to numerous results in topology and algebraic geometry; projective algebra, including the elements of projective geometry; and differential algebra, where general algebra yielded direct results in the theory of differential equations. The theory of semigroups and that of quasigroups ceased to be simply theories of 'generalized' groups and found their own paths of development and their own areas of application. Eventually, the general theory of universal algebras came into being, as well as the even more general theory of models, which is interwoven with mathematical logic.

One should have thought that the fundamental ideas and the most important results accumulated in present-day general algebra ought to be part of the scientific equipment of every well-educated mathematician to the same extent as in the thirties, when the majority of candidates in mathematics were examined in modern algebra. However, this is by no means the case: a wide circle of mathematicians today has an acquaintance with the achievements of general algebra that remains rather on the level of the early thirties.

The reason is not difficult to see. The basic textbook from which young [Russian] mathematicians learn their algebra is still van der Waerden's book. Although this book undoubtedly had a most remarkable and outstanding role in the history of mathematics in the twentieth century, it is now so far removed from the present state of algebra that the author himself changed its title, in the fourth edition, simply to *Algebra*.

In the non-Russian mathematical literature there are other, more recent books. Some of them, by modernizing the material in van der Waerden's book somewhat, essentially supplement and develop it in the direction of the personal scientific interest of their authors. They are valuable books, but they do not give a true picture of the present state of general algebra. Besides, these books are, as a rule, of considerable bulk and are addressed to algebraists rather than to mathematicians regardless of their specialization. Books of another type present essentially a summary of the most fundamental algebraic concepts and their simplest properties. Although useful as reference books, they do not enable the reader to appreciate the full generality and depth of modern algebraic research: the deepest and most significant results are either omitted entirely or are stated as exercises.

In order to show mathematicians the new look of general algebra, a book of another character is required. Not very large in size, it would have to be addressed to a reader who, having mastered a university course in higher algebra, wishes to extend his algebraic education without necessarily intending to choose algebra as his field of specialization. Of course,

this is not to rule out the possibility that even an algebraist, in regard to problems at some remove from his particular interests, might find something useful in the book.

Such a book should not, and could not, replace monographs on the various branches of general algebra. Nor should it merely string together introductory chapters to such monographs. The object of the book would be to exhibit the main branches of modern general algebra, preferably in their mutual interconnection, the exposition being restricted to individual important theorems and aiming straight at these theorems. The selection of a fairly small number of such theorems in each of the main branches of general algebra would inevitably be influenced by the personal judgment of the author of the book. The theorems themselves would by no means have to be given in the greatest generality now attainable.

The contents of the book would, of course, be rather like a mosaic; and, in following the author, the reader would on occasion have to pass, within the confines of a single section, from one branch of general algebra to another. The division of the material into chapters would be so provisory that there could be no question of a scheme of interdependence of the chapters.

At the All-Soviet Congress on Algebra and Number Theory in 1951 (see *Usp. Math. Nauk* (1952), vol. 7, part 3, page 167) I spoke of the desirability of a book of this character, and I began to write it in 1956. During the four years that have since passed, work on the book was repeatedly interrupted and then resumed; several sections were written and rewritten; finished material was rearranged, altered, discarded.... In other words, the work assumed such a character that more and more often I was reminded of Balzac's story *Le chef-d'oeuvre inconnu*. So it was reasonable to complete the work without getting the book into a form that would correspond to the program laid down above.

The title of the book is entirely justified by the fact that it is based on three extensive special courses I have given at the University of Moscow during the last ten years.

Here and there in the book are statements of certain results that are neither proved nor used. It is hoped that the reader will not skip these passages, which are set apart from the body of the text by the symbol ❖. It goes almost without saying that the inclusion of such supplementary references does not mean that the corresponding sections of the book necessarily represent the most recent results.

When an article in a periodical is quoted in the book, this is rather an incidental matter and should not be taken as material on the history of algebra in the twentieth century. On the other hand, a fairly complete

list of books on various branches of general algebra that have appeared in the last thirty years is included in the book. This list contains some survey papers as well.

Inasmuch as the book embodies many plans in its mosaic-like structure, there are very frequent references to preceding material, although in the majority of cases the reader will no doubt find these references superfluous.

I have had the pleasure of presenting the original plan of the book, as well as a number of chapters, some of them in various drafts, before the Seminar on General Algebra at the University of Moscow. To those who participated in the Seminar I offer my sincere thanks for their interest in my work and for their advice and criticism. I also wish to express my warmest thanks to Oleg N. Golovin, who has taken on the heavy burden of editing the book, for the great care with which he has read the manuscript and the many helpful suggestions he has made.

Moscow, May 1960 A. G. KUROSH

CONTENTS

AUTHOR'S PREFACE TO ENGLISH EDITION 3

INTRODUCTION 5

I. RELATIONS 13

§ 1. Sets 13
§ 2. Binary Relations 15
§ 3. Equivalence Relations 18
§ 4. Partial Order 20
§ 5. Minimal Condition 22
§ 6. Theorems Equivalent to the Axiom of Choice 26

II. GROUPS AND RINGS 30

§ 7. Groupoids, Semigroups, and Groups 30
§ 8. Rings and Fields 35
§ 9. Subgroups and Subrings 42
§ 10. Isomorphism 46
§ 11. Embedding of Semigroups in Groups and of Rings in Fields 51
§ 12. Non-Associative Fields, Quasigroups. Isotopy 58
§ 13. Normal Subgroups, Ideals 63
§ 14. Gaussian Semigroups 71
§ 15. Gaussian Rings 77
§ 16. Dedekind Rings 84

III. UNIVERSAL ALGEBRAS. GROUPS WITH MULTI-OPERATORS .. 93

§ 17. Universal Algebras, Homomorphisms 93
§ 18. Groups with Multi-operators 98
§ 19. Automorphisms and Endomorphisms. The Field of p-adic Numbers 108

§ 20. Normal Series and Composition Series 118
§ 21. Abelian, Nilpotent, and Solvable Ω-groups 123
§ 22. Primitive Classes of Universal Algebras 131
§ 23. Free Universal Algebras 134
§ 24. Free Products of Groups 143

IV. LATTICES 155

§ 25. Lattices and Complete Lattices 155
§ 26. Modular Lattices 163
§ 27. Direct Unions. The Schmidt-Ore Theorem 169
§ 28. Direct Decompositions of Ω-groups 178
§ 29. Complete Direct Sums of Universal Algebras 182
§ 30. Distributive Lattices 186

V. OPERATOR GROUPS AND OPERATOR RINGS. MODULES. LINEAR ALGEBRAS 191

§ 31. Operator Groups and Operator Rings 191
§ 32. Free Modules. Abelian Groups 198
§ 33. Vector Spaces Over Fields 205
§ 34. Rings of Linear Transformations 209
§ 35. Simple Rings. Jacobson's Theorem 214
§ 36. Linear Algebras. The Quaternion Algebra and the Cayley Algebra 221
§ 37. Alternative Rings. Artin's Theorem 229
§ 38. The Generalized Frobenius Theorem 234
§ 39. The Birkhoff-Witt Theorem on Lie Algebras 243
§ 40. Derivations. Differential Rings 248

VI. ORDERED AND TOPOLOGICAL GROUPS AND RINGS. RINGS WITH A VALUATION 255

§ 41. Ordered Groups 255
§ 42. Ordered Rings 261
§ 43. Archimedean Groups and Rings 266
§ 44. Rings with a Valuation 273
§ 45. Logarithmic Valuations of Fields 278

§ 46. Albert's Theorem on Normed Algebras 283
§ 47. Closure. Topological Spaces 289
§ 48. Special Types of Topological Spaces 296
§ 49. Topological Groups 299
§ 50. Connection Between Topology and Valuations in Rings and Fields 304
§ 51. Galois Connections. The Fundamental Theorem of Galois Theory 313

NOTE ON TERMINOLOGY 321

BIBLIOGRAPHY 322

INDEX 327

CHAPTER I

RELATIONS

§ 1. Sets

1. Basic to General Algebra are the concepts and methods of the theory of sets. The reader who begins the study of General Algebra is, of course, familiar with these, so that he need only be reminded of the definitions of such set-theoretical concepts as *subset, complement* of a subset in a set, the *empty set,* and *intersection* and *union* of sets. We shall adopt the following notation: For the intersection and union of sets we shall use the symbols $\cap$ and $\cup$, respectively; to express the fact that a subset or an element is contained in a set, we use the symbols $\subset$ and $\in$; and the complement of a subset A in a set M will be denoted by $M \setminus A$.

The operations of intersection and union of sets are connected by the following distributive laws, which are dual to each other: *For any three sets A, B, and C,*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \tag{1}$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \tag{2}$$

Let us prove the second of these identities. Since $B \cap C \subseteq B$, we have

$$A \cup (B \cap C) \subseteq A \cup B;$$

similarly,

$$A \cup (B \cap C) \subseteq A \cup C,$$

and so the left-hand side of (2) is contained in the right-hand side. On the other hand, if an element x is contained in the right-hand side of (2), then we have simultaneously

$$x \in (A \cup B) \quad \text{and} \quad x \in (A \cup C). \tag{3}$$

If $x \in A$, then x is contained in the left-hand side of (2). But if x does not belong to A, then it follows from (3) that x belongs both to B and to C, in other words, that it is in the intersection $B \cap C$, so that again x is contained

in the left-hand side of (2). Therefore, the whole right-hand side of (2) is included in the left-hand side, and equation (2) is proved. // [1]

2. The reader is also familiar with the concept of a *mapping*, or *correspondence*, or *single-valued function*. To indicate that φ is a mapping of a set A into a set B, i.e., onto the whole of B or onto a subset of B, we shall use the symbol $\varphi : A \to B$, and we shall denote the image of an element $a \in A$ under the mapping φ by $a\varphi$.

If $\varphi : A \to B$ and $\psi : B \to C$, then by carrying out the mappings φ and ψ in succession we obtain a well-defined mapping of A into C, which we denote by $\varphi\psi$ and refer to as the *product* of φ and ψ. Thus, for all a in A,

$$a(\varphi\psi) = (a\varphi)\psi. \tag{4}$$

This multiplication of mappings can be called *partial*: if two arbitrary mappings $\varphi : A \to B$ and $\psi : A' \to B'$ are given, then the product $\varphi\psi$ does not always exist; it exists if and only if we have $a\varphi \in A'$ for all $a \in A$. Hence it follows that *for mappings of an arbitrary set A into itself the product always exists.*

Multiplication of mappings is associative: *if three mappings are given,*

$$\varphi : A \to B, \quad \psi : B \to C, \quad \chi : C \to D, \tag{5}$$

then

$$(\varphi\psi)\chi = \varphi(\psi\chi). \tag{6}$$

For if a is an arbitrary element of A, then by (4),

$$a[(\varphi\psi)\chi] = [a(\varphi\psi)]\chi = [(a\varphi)\psi]\chi = (a\varphi)(\psi\chi) = a[\varphi(\psi\chi)]. \; //$$

In virtue of (6), the result of carrying out the mappings (5) in succession can be written as $\varphi\psi\chi$.

3. The *identity mapping* of a set A into itself will be denoted by ε_A; thus:

$$a\varepsilon_A = a \quad \text{for all} \quad a \in A.$$

The identity mapping plays the role of the unit element in the multiplication of mappings, since for arbitrary mappings $\varphi : A \to B$ and $\psi : C \to A$,

$$\varepsilon_A\varphi = \varphi \quad \text{and} \quad \psi\varepsilon_A = \psi.$$

The reader is familiar with the concept of a *one-to-one mapping* of a set A onto a set B (i.e., a one-to-one correspondence between these sets). Clearly, the mapping $\varphi : A \to B$ is a one-to-one mapping of A onto B if and only if

[1] This symbol signifies the conclusion of a proof.

it has an inverse mapping, i.e., a mapping $\varphi^{-1}\varphi\colon B \to A$ satisfying the conditions

$$\varphi\varphi^{-1} = \varepsilon_A \quad \text{and} \quad \varphi^{-1}\varphi = \varepsilon_B.$$

If there exists a one-to-one mapping of a set A onto a set B, then the sets A and B are said to be *equipotent* or to have the same *cardinal number*, or *power*. The cardinal number of a finite set coincides with the number of its elements; sets that are equipotent with the set of all natural numbers are called *countable*; and sets that are equipotent with the set of all real numbers are said to have the *cardinal number of the continuum*.

4. In the study of infinite sets it is very often necessary to use the following *Axiom of Choice*:

If M is any set, then there exists a function that assigns to every non-empty subset A of M a definite element $\varphi(A)$ of that subset.

In other words, *the function φ selects from* each non-empty subset of M exactly one element of that subset.

What is the logical basis of this axiom? Are we justified in applying it? These are very difficult and controversial questions in the foundations of the theory of sets. However, we cannot get along without this axiom. For countable sets, though, it can easily be proved: if the elements of a set M are enumerated by the natural numbers, then we obtain the required function by selecting in each subset A of M that element which has the smallest number.

In § 6.3 we shall give a number of statements that are equivalent to the Axiom of Choice.

§ 2. Binary Relations

1. If M is a set, then its *square* $M \times M$ is defined as the set of all ordered pairs (a, b), where $a, b \in M$. Let R be an arbitrary subset of $M \times M$. In the following way R defines a *binary relation* on M (the symbol R will also be used to denote the relation itself; but in concrete cases, various special symbols are used to write down relations): if $a, b \in M$, then we say that the *element a has the relation R to the element b* and we express this by $a\,R\,b$ if and only if the pair (a, b) belongs to the subset R; in other words, the notations

$$a\,R\,b \qquad \text{and} \qquad (a, b) \in R$$

are equivalent.

The study of binary relations on a set M does not differ, therefore, from the study of subsets of $M \times M$. In particular, we can speak of the *inclusion* of a binary relation R in a binary relation R', $R \subseteq R'$, and of the *intersec-*

tion and *union* of binary relations. The *complement* to a binary relation R is the binary relation $\bar{R}$ defined by the subset $\bar{R} = (M \times M) \setminus R$; in other words, $a \bar{R} b$ if and only if $(a, b) \notin R$.

2. On the other hand, the fact that binary relations are given by sets of ordered pairs of elements of M makes the algebra of binary relations richer than the simple algebra of subsets of an arbitrary set. Suppose, for example, that arbitrary binary relations R and S are given on a set M. Their *product* RS is the binary relation that is defined as follows: $a(RS)b$ if and only if at least one element c exists in M such that aRc and cSb.

Multiplication of binary relations is associative,

$$(RS)T = R(ST),$$

since an element a has either of the relations $(RS)T$ and $R(ST)$ to the element b if and only if there exist elements c and d such that

$$aRc, \quad cSd, \quad dTb.$$

However, multiplication of binary relations is not commutative; only occasionally are binary relations R and S *permutable*:

$$RS = SR.$$

If binary relations R_i (where i ranges over an index set I) and S are given on a set M, then

$$\Big(\bigcup_{i \in I} R_i\Big)S = \bigcup_{i \in I} R_i S, \qquad S\Big(\bigcup_{i \in I} R_i\Big) = \bigcup_{i \in I} S R_i. \tag{1}$$

For $a\Big[\Big(\bigcup_{i \in I} R_i\Big) S\Big] b$ is equivalent to the existence of an element c such that $a\Big(\bigcup_{i \in I} R_i\Big) c$ and cSb. This is equivalent, however, to the existence of an index i_0 such that $aR_{i_0}c$ and cSb, i.e., $a(R_{i_0}S)b$, and therefore $a\Big(\bigcup_{i \in I} R_i S\Big) b$. //

Note that in the equations (1) union cannot be replaced by intersection.

It follows from (1) that *if binary relations R, R', and S are given, where $R \subseteq R'$, then*

$$RS \subseteq R'S \quad \textit{and} \quad SR \subseteq SR'. \tag{2}$$

For the inclusion $R \subseteq R'$ is equivalent to $R \cup R' = R'$, and this implies the equation

$$(R \cup R')S = RS \cup R'S = R'S,$$

which is equivalent to the inclusion $RS \subseteq R'S$. //

3. For every binary relation R on a set M there exists the *inverse relation* R^{-1}, which is defined as follows: $aR^{-1}b$ if and only if bRa. It is clear that

$$(R^{-1})^{-1}=R$$

and that $R\subseteq S$ implies that $R^{-1}\subseteq S^{-1}$.

If binary relations R_i $(i\in I)$, S, and T are given on a set M, then

$$\Big(\bigcap_{i\in I}R_i\Big)^{-1}=\bigcap_{i\in I}R_i^{-1}, \tag{3}$$

$$\Big(\bigcup_{i\in I}R_i\Big)^{-1}=\bigcup_{i\in I}R_i^{-1}, \tag{4}$$

$$(ST)^{-1}=T^{-1}S^{-1}. \tag{5}$$

For $a\Big(\bigcap_{i\in I}R_i\Big)^{-1}b$ means that $b\Big(\bigcap_{i\in I}R_i\Big)a$, i.e., bR_ia for all $i\in I$, hence $aR_i{}^{-1}b$ for all $i\in I$, and therefore $a\Big(\bigcap_{i\in I}R_i{}^{-1}\Big)b$. (4) is proved similarly. Finally, $a(ST)^{-1}b$ means that $b(ST)a$, i.e., that there exists an element c such that bSc and cTa, so that $aT^{-1}c$ and $cS^{-1}b$, or $a(T^{-1}S^{-1})b$. //

4. The *identity relation* E is defined as follows: aEb if and only if $a=b$; in other words, the relation E is given by the set of all pairs of the form (a,a), where $a\in M$. Obviously, $E^{-1}=E$, and for every binary relation R,

$$RE=ER=R.$$

We also mention the *empty relation* O, which is defined by the empty subset of $M\times M$. It is clear that for every binary relation R on M

$$O\subseteq R \quad\text{and}\quad RO=OR=O.$$

5. In the next few sections we shall be concerned with binary relations R given on a set M that have some of the following four properties:

Reflexivity: aRa for all $a\in M$; in other words, $E\subseteq R$.
Transitivity: If aRb and bRc, then aRc; in other words, $RR\subseteq R$.
Symmetry: If aRb, then bRa; in other words, $R^{-1}=R$.
Antisymmetry: If aRb and bRa, then $a=b$; in other words,

$$R\cap R^{-1}\subseteq E.$$

If a binary relation R has any one of these four properties, then the inverse relation R^{-1} has the same property.

For if $E\subseteq R$, then

$$E=E^{-1}\subseteq R^{-1}$$

If $RR \subseteq R$, then

$$R^{-1}R^{-1} = (RR)^{-1} \subseteq R^{-1}.$$

If $R^{-1} = R$, then

$$(R^{-1})^{-1} = R = R^{-1}.$$

Finally, if $R \cap R^{-1} \subseteq E$, then

$$R^{-1} \cap (R^{-1})^{-1} = R^{-1} \cap R \subseteq E. /\!/$$

6. Let N be a subset of a set M and R a relation on M. There is a natural way of *inducing* a binary relation R^N in the set N: if $a, b \in N$, then aR^Nb if and only if aRb holds in M. In other words, taking into account that $N \times N$ is a subset of $M \times M$,

$$R^N = R \cap (N \times N).$$

It is easy to verify the following equations:

$$(R^N)^{-1} = (R^{-1})^N, \qquad \Big(\bigcap_{i \in I} R_i\Big)^N = \bigcap_{i \in I} R_i^N, \qquad \Big(\bigcup_{i \in I} R_i\Big)^N = \bigcup_{i \in I} R_i^N.$$

7. The concept of a binary relation is susceptible of various generalizations. For example, let us consider the *n-th power* M^n of a set M, that is, the set of all ordered systems $(a_1, a_2, \ldots, a_n)$ of n elements of M. Then every subset R of M^n defines an *n-ary relation* on M (for $n = 3$, a *ternary relation*). Sets on which a number of such relations are given are called *models* and are the object of an independent theory.

§ 3. Equivalence Relations

1. An important type of binary relation is that of the *equivalence relations*, that is, binary relations having the properties of *reflexivity*, *transitivity*, and *symmetry* (see § 2.5). Of the numerous concrete examples of such relations known to the reader we only mention equality of fractions and congruence of integers with respect to some modulus. To denote equivalence relations we shall for the most part use the symbols $\sim$ and $\equiv$.

2. Equivalence relations defined on a set M are very closely connected with *partitions of M into disjoint classes*. By a partition we mean a choice of a system of non-empty subsets of M (the *classes* of this partition) such that every element of M belongs to one and only one of these subsets.

Every partition π of a set M defines an equivalence relation on M.

For if $a, b \in M$ and if we set $a \sim b$ if and only if a and b belong to one and the same class of the partition π, then we obtain a binary relation on M

which obviously satisfies all the requirements of the definition of an equivalence relation. //

Conversely, *every equivalence relation R given on a set M defines a partition of this set.*

For let us define the *class* of an element a as the set of all those elements x of M for which aRx, and let us denote this class by K_a. From the reflexivity of R it follows that $a \in K_a$, so that the system of classes K_a $(a \in M)$ covers the whole set M. Further, the symmetry of R shows that it follows from $b \in K_a$ that $a \in K_b$, and the transitivity of R ensures that if $b \in K_a$, then it follows from $c \in K_b$ that $c \in K_a$, i.e., that $K_b \subseteq K_a$. These last remarks, taken together, lead to the fact that if $b \in K_a$, then $K_b = K_a$, i.e., that a class is determined by any one of its elements. Finally, if K_a and K_b are two arbitrary classes with a non-empty intersection containing, say, the element c, then $K_a = K_c$ and $K_b = K_c$, i.e., the classes K_a and K_b coincide. We have thus shown that the system of all distinct classes of the form K_a is a partition of M. //

It is obvious that the transition from a partition π of M to the equivalence relation defined by it, and then to the partition of M defined by this equivalence relation leads back to the partition π. *There is a one-to-one correspondence between the equivalence relations on a set M and the partitions of M into disjoint classes.*

3. *If equivalence relations R_i $(i \in I)$ are given on a set M, then their intersection is also an equivalence relation.*

For from aR_ia for all $i \in I$ it follows that $a\left(\bigcap_{i \in I} R_i\right)a$. Next, if $a\left(\bigcap_{i \in I} R_i\right)b$ and $b\left(\bigcap_{i \in I} R_i\right)c$, then aR_ib and bR_ic for all $i \in I$, so that aR_ic for all $i \in I$, and therefore $a\left(\bigcap_{i \in I} R_i\right)c$. Finally, if $a\left(\bigcap_{i \in I} R_i\right)b$, then aR_ib for all $i \in I$, that is, bR_ia for all $i \in I$, and hence $b\left(\bigcap_{i \in I} R_i\right)a$. //

It is not difficult to verify that when equivalence relations R_i $(i \in I)$ correspond to partitions π_i of M $(i \in I)$, then the equivalence relation $\bigcap_{i \in I} R_i$ corresponds to the partition whose classes are all the non-empty intersections of the classes taken for each of the partitions π_i $(i \in I)$ in turn. We call this partition the *intersection* of the partitions π_i $(i \in I)$.

If equivalence relations R_i $(i \in I)$ are given on a set M, then their union, interpreted in the sense of the union of binary relations, need not, in general, be an equivalence relation. However, *there exists an equivalence*

relation on M that includes all the relations R_i $(i \in I)$ (in the sense of inclusion of binary relations) *and is itself included in every other equivalence relation that includes all the R_i $(i \in I)$.* This equivalence relation can be called the *union* of the equivalence relations R_i $(i \in I)$.

To prove this, we define a binary relation S on M as follows: aSb if and only if there is at least one way of choosing a finite system of elements of M

$$a = c_0, c_1, c_2, \ldots, c_{n-1}, c_n = b, \tag{1}$$

such that for $k = 1, 2, \ldots, n$ there exists at least one index $i_k \in I$ for which $c_{k-1} R_{i_k} c_k$. The relation S is obviously reflexive, transitive, and symmetric. If T is any equivalence relation that includes all the R_i $(i \in I)$ and if aSb, with the corresponding system of elements (1), then it follows from $c_{k-1} R_{i_k} c_k$ that $c_{k-1} T c_k$ $(k = 1, 2, \ldots, n)$, so that by the transitivity of T we have aTb. //

❖ The product RS of two equivalence relations R and S is an equivalence relation if and only if R and S are permutable: $RS = SR$. If this is the case, then the union of R and S coincides with their product as binary relations [F. Šik, *Spisy vyd. přírodověde fakult. Masarykovy univ.* (1954), No. 3, pp. 97-102]. ❖

4. The set of classes of the partition corresponding to a given equivalence relation R on a set M will be denoted by M/R and called the *factor set* of M with respect to R. The mapping of M onto M/R which assigns to every element $a \in M$ that class of the partition corresponding to R which contains a is called the *natural mapping* of M onto M/R.

Between the equivalence relations on a set M and the mappings of this set onto other sets there is a close connection, the prototype of the so-called *homomorphism theorems* that we shall encounter a number of times in the following chapters. For a mapping φ of M onto some other set N gives rise to a well-defined equivalence relation R on M (i.e., a partition of this set): we put aRb for elements $a, b \in M$ if and only if $a\varphi = b\varphi$. By assigning to each element x of N the class of those elements of M that have x as their image under the mapping φ *we obtain a one-to-one mapping ξ of N onto M/R for which the product $\varphi\xi$ coincides with the natural mapping of M onto M/R.*

§ 4. Partial Order

1. Another very important type of binary relation is constituted by the *relations of partial order,* that is, binary relations having the properties of *reflexivity, transitivity,* and *antisymmetry.* A set M on which a partial

order is given is called *partially ordered*. To denote partial order we shall use the symbol $\leqq$; if $a, b \in M$ and $a \leqq b$, then depending on the circumstances we say that *a is less than or equal to b, a is contained in b, a precedes b*.

If $a \leqq b$ and $a \neq b$, we write $a < b$ and say that *a is less than b, a is strictly contained in b*, etc. The binary relation $<$, of course, is not reflexive. We shall denote by $\geqq$ and $>$ the relations that are inverse to the relations $\leqq$ and $<$; for example: $a \geqq b$ (*a is greater than or equal to b, a contains b, a follows b*) if and only if $b \leqq a$.

Let M be a set on which a partial order is given. The elements a and b of this set are called *comparable* if $a \leqq b$ or $b \leqq a$. It is by no means true that any two elements of M must necessarily be comparable—this is precisely the reason why we speak of a 'partial' order. Thus, we obtain a *trivial* partial order of a set M by setting $a \leqq b$ only for $a = b$; distinct elements of M are in this case incomparable. A partially ordered set in which any two elements are comparable is called an *ordered set*, or *linearly ordered set*, or *chain*.

2. Ordered and partially ordered sets occur extremely often in various branches of mathematics. As our first few examples of ordered sets we mention the set of natural numbers and the set of points on a straight line (i.e., the set of all real numbers), both of them in their natural order. Some examples of partially (but not linearly) ordered sets are the following:

The set $\tilde{N}$ of all subsets of a given set N, with the relation of set-theoretical inclusion $\subseteq$ as the relation of partial order;

The set of all continuous real functions defined on the interval $[0, 1]$, if $f \leqq g$ means that $f(x) \leqq g(x)$ for all x in this interval.

The set of all natural numbers, if $a \leqq b$ is interpreted to mean that b is divisible by a.

❖ Every partial order of a given set M can be extended to a linear order of the same set, i.e., can be included in a linear order (in the sense of inclusion of binary relations; see § 2.1) [Shpilrain, Fund. Math., vol. 16 (1930), pp. 386-389]. ❖

3. Let φ be a one-to-one correspondence between the partially ordered sets M and M'

$$a\varphi = a' \quad (a \in M;\ a' \in M').$$

If $a \leqq b$, where $a, b \in M$, always implies that $a\varphi \leqq b\varphi$, and vice versa, then φ is called an *isomorphism* between M and M', and M and M' are called *isomorphic* partially ordered sets. Obviously, in all cases where the partial order is the primary object of study and the nature of the elements constituting the sets in question does not play any role, isomorphic sets can be regarded as identical.

4. We know (see § 2.6) that a partial order on a set M induces a certain binary relation on every subset; it is easy to see that this is also one of partial order. We shall say that a partially ordered set M is *isomorphically embedded* in a partially ordered set N if there exists an isomorphic mapping of M onto a subset N' of N, where N' is furnished with the partial order induced in it by the partial order of N.

5. The following theorem emphasizes the fundamental role of the first example of a partially ordered set given in § 4.2.

Every partially ordered set M can be isomorphically embedded in the set $\tilde{N}$ of all subsets of a certain set N, partially ordered by inclusion. We can take M itself, for example, as our N.

In fact, let us assign to every element a of M the subset A consisting of those elements $x \in M$ for which $x \leqq a$. Let $a, b \in M$ and let A, B be the subsets corresponding to them. If $A = B$, then $b \leqq a$ and $a \leqq b$, and hence $a = b$. This shows that the correspondence $a \to A$ is a one-to-one mapping of M into the set $\tilde{M}$ of all its subsets. Furthermore, if $a \leqq b$, then it follows from $x \leqq a$ that $x \leqq b$, so that $A \subseteq B$. Conversely, if $A \subseteq B$, then $a \in B$, so that $a \leqq b$. Thus, the correspondence $a \to A$ is an isomorphic embedding of M in $\tilde{M}$. //

6. From what was shown in § 2.5 it follows that *the relation inverse to a relation of partial order is itself a partial order.*

Partially ordered sets M and M' are called *inversely isomorphic* if each is isomorphic to the other taken with the inverse partial order, that is, if there exists between them a one-to-one correspondence φ,

$$a\varphi = a' \qquad (a \in M\,;\ a' \in M')$$

such that, for $a \leqq b$ $(a, b \in M)$ if and only if $a\varphi \geqq b\varphi$.

§ 5. Minimal Condition

1. An element a of a partially ordered set M is called a *minimal element* of the set if there is no element x in M with $x < a$. Clearly, M may contain many distinct minimal elements, and it need not contain any.

Thus, the set $\tilde{N}$ of all subsets of a set N has a unique minimal element—the empty subset. In the set of all non-empty subsets of N, the minimal elements are all the subsets consisting of a single element. Finally, if the set N is infinite, then the set of all its infinite subsets has no minimal elements at all.

The concept of a minimal element will now be utilized in defining a special class of partially ordered sets wider than the class of finite partially ordered

sets. This is the class of partially ordered sets satisfying any of the following three equivalent conditions:

MINIMAL CONDITION: *Every non-empty subset N of the partially ordered set M has at least one element that is minimal (in N).*

DESCENDING CHAIN CONDITION:[1] *Every strictly descending chain of elements of the partially ordered set M*

$$a_1 > a_2 > \ldots > a_n > \ldots$$

terminates for some finite value of n. In other words, for every descending chain of elements

$$a_1 \geqq a_2 \geqq \ldots \geqq a_n \geqq \ldots$$

there is an index n at which the chain becomes stationary, i.e.,

$$a_n = a_{n+1} = \ldots .$$

CONDITION OF INDUCTIVENESS: *All the elements of a partially ordered set M have a certain property $\mathfrak{E}$ if all the minimal elements of the set (provided they exist) have this property and if from the fact that $\mathfrak{E}$ holds for all elements strictly preceding a certain element a it can be deduced that the property holds for a itself.*

2. Let us prove the equivalence of these three conditions.

The condition of inductiveness follows from the minimal condition.

For suppose that a partially ordered set M satisfies the minimal condition and that for a certain property $\mathfrak{E}$ the premises of the condition of inductiveness hold in M. If, nevertheless, M contains elements that do not have the property $\mathfrak{E}$, then let a be a minimal one among these elements: the existence of a follows from the minimal condition. The element a cannot be minimal in the whole of M, as follows from the first premise of the condition of inductiveness; and since all the elements strictly preceding a have the property $\mathfrak{E}$, by the second premise of the condition of inductiveness a itself must have the property $\mathfrak{E}$, and we have a contradiction. //

The descending chain condition follows from the condition of inductiveness.

For suppose that a partially ordered set M satisfies the condition of inductiveness. We apply this condition to the following property: An element a will be said to have the property $\mathfrak{E}$ if every strictly descending chain of elements beginning with a breaks off—in other words, if it is

[1] The concept of a descending chain that occurs in this condition is a special case of the general concept of a chain introduced in § 4.1.

necessarily a finite chain. All the minimal elements of M, if they exist, obviously have this property. Moreover, suppose that all the elements strictly preceding a have our property $\mathfrak{E}$. In that case the second term of every strictly decreasing chain beginning with a has the property $\mathfrak{E}$, and therefore the chain in question must break off, i.e., a itself also has the property $\mathfrak{E}$. From the condition of inductiveness it now follows that all the elements of M have our property $\mathfrak{E}$, i.e., that every strictly descending chain in M breaks off—it must begin, of course, with a certain element. //

The minimal condition follows from the descending chain condition.

To prove this, we assume that a partially ordered set M does not satisfy the minimal condition, that is, that a non-empty subset N of M has no minimal elements. Using the Axiom of Choice (see § 1.4) we select just one element in every non-empty subset of N and then we construct a sequence of elements

$$a_n \qquad (n = 1, 2, \ldots) \tag{1}$$

in the following way. For a_1 we choose the element selected in N itself. If a_n is already determined and $a_n \in N$, then we take for a_{n+1} the element selected in the non-empty set of elements of N strictly preceding a_n—the set is non-empty because N has no minimal elements. The sequence (1) is obviously an infinite strictly descending chain, so that the set N cannot satisfy the descending chain condition. //

3. The condition of inductiveness enables us to carry out not only proofs by induction, but also *constructions by induction.* Suppose that M is a partially ordered set with minimal condition and that we wish to define on this set a function $\varphi(x)$ that assigns to every element x of M a certain element of an auxiliary set S. We shall assume here that $\varphi(x)$ must satisfy a certain *recurrence relation,* i.e., a relation that for every $a \in M$ determines the value $\varphi(a)$ uniquely from the values $\varphi(b)$ for all b strictly less than a. We shall show that *there exists one and only one function $\varphi(x)$, defined on the whole set M, satisfying the given recurrence relation and assuming arbitrary preassigned values on all the minimal elements of M.*

Let us begin with the proof of the *uniqueness.* Suppose that on M there exist two distinct functions, $\varphi(x)$ and $\psi(x)$, satisfying our conditions. By the minimal condition, the non-empty set of those elements x for which $\varphi(x) \neq \psi(x)$ has at least one element a. This element cannot be minimal in the whole of M, because the functions $\varphi(x)$ and $\psi(x)$ by hypothesis coincide on the minimal elements of M. Therefore, there exist elements b with $b < a$; and for all these elements, $\varphi(b) = \psi(b)$. However, the recurrence relation uniquely determines the value of the functions in question for

$x = a$ from their values for all $b < a$, and therefore $\varphi(a) = \psi(a)$, so that we have arrived at a contradiction.

Now we proceed to the proof of the *existence* of the required function $\varphi(x)$ under the assumption that its values on the minimal elements are given. We shall say that an element $a \in M$ has the property $\mathfrak{E}$ if on the set A of all x with $x \leqq a$ a function $\varphi_a(x)$ can be defined satisfying the given recurrence relation and assuming the given values on the minimal elements of M contained in A.

All the minimal elements of M obviously have the property $\mathfrak{E}$. On the other hand, if a and b have the property $\mathfrak{E}$ and $b < a$, then by applying the uniqueness of the required function, proved above, not to the set M, but to the set B of those x for which $x \leqq b$, we see that for all such x

$$\varphi_b(x) = \varphi_a(x).$$

Hence it follows that if all the elements b strictly preceding a given element a have the property $\mathfrak{E}$, then a itself also has this property; we obtain the function $\varphi_a(x)$ satisfying all the conditions by putting for every b $(b < a)$

$$\varphi_a(b) = \varphi_b(b),$$

and by choosing for $\varphi_a(a)$ the value that is uniquely determined by the recurrence relation.

By the condition of inductiveness, which holds in M, we can now claim that all the elements of this set have the property $\mathfrak{E}$. Finally, by setting

$$\varphi(a) = \varphi_a(a)$$

for all $a \in M$, we define the function $\varphi(x)$ having all the required properties. //

4. A linearly ordered set satisfying the minimal condition, and therefore the other two conditions equivalent with it, is called *well-ordered*. An example of a well-ordered set is the set of natural numbers in its natural order. Every subset of a well-ordered set is itself well-ordered. From the definition of a well-ordered set it follows that it has a *unique* minimal element.

A well-ordered set contains, for every element a, an element immediately following upon a. However, a need not have an immediate predecessor; in that case it is called a *limit element*.

A partially ordered set satisfies the minimal condition if and only if all its chains (i.e., linearly ordered subsets) are well-ordered.

For if a partially ordered set M satisfies the minimal condition, then this is also true of all its subsets and, in particular, of all its chains. The con-

verse assertion follows from the fact that in the statement of the descending chain condition—which is equivalent to the minimal condition—only the chains of M are used. //

5. In a partially ordered set M we can go over to the inverse partial order. The minimal elements of this inverse order are called *maximal elements* of M in its original order; and descending chains in the inverse order are called *ascending chains* of M. Quite generally, for every concept (or every statement) connected with a partial order we can obtain in this way a *dual* concept (or statement).

Let M be a partially ordered set satisfying the minimal condition. By taking the inverse partial order in M we obtain a partially ordered set satisfying the *maximal condition*. All that has been said above concerning sets with minimal condition remains true for sets with maximal condition after replacement of the relation $\leqq$ by $\geqq$ and vice versa.

§ 6. Theorems Equivalent to the Axiom of Choice

1. Let N be a subset of a partially ordered set M. Every element a of M (not necessarily contained in N) that satisfies the condition $a \geqq x$ for all $x \in N$ is called an *upper bound* of N in M. The concept of a *lower bound* is dual to this.

2. Next, if M is a partially ordered set, then the set of all its chains can be partially ordered by set-theoretical inclusion. The maximal elements of the latter set, provided they exist, are naturally called *maximal chains* of M.

3. These concepts are used in the statement of two of the following three theorems *each of which*, as will be shown presently, *is equivalent to the Axiom of Choice* (see § 1.4).

ZERMELO'S THEOREM: *Every set can be well-ordered.*

HAUSDORFF'S THEOREM: *Every chain of a partially ordered set is contained in a maximal chain.*

THE KURATOWSKI-ZORN THEOREM: *If every chain of a partially ordered set M has an upper bound, then every element of M is contained in a maximal element.*

4. Let us now show the equivalence of these theorems and the Axiom of Choice.

Zermelo's Theorem follows from the Axiom of Choice.

We define a *segment* of a well-ordered set A to be a subset B of A such that if B contains b it also contains all the $x \in A$ for which $x \leqq b$. The set of elements that strictly precede some element a of A is a *proper* segment of

A (i.e., a segment different from A itself) and these are the only proper segments: if B is such a segment, then B consists of all the elements that strictly precede the minimal element of the complement $A \setminus B$, so that it is *determined* by this minimal element. We shall also count the empty subset among the proper segments of A; it is determined by the minimal element of the set.

We now proceed to the proof of the theorem. Let M be an arbitrary set. On the basis of the Axiom of Choice, we select just one element $\varphi(N)$ in every non-empty subset N of M. We shall call a non-empty subset A of M *distinguished* if it can be well-ordered in such a way that for every $a \in A$

$$a = \varphi(M \setminus A'),$$

where A' is the segment of A in the given well-ordering determined by a. M has distinguished subsets: for example, the subset consisting of the single element $\varphi(M)$.

Let A and B be two distinguished subsets for which well-orderings with the property mentioned in the preceding paragraph have been chosen. Then both these subsets have $\varphi(M)$ as their first element and therefore they have non-empty segments in common. The union C of all common segments of these two subsets is obviously a segment in each of them; it is the largest among the common segments. If C were different from A as well as from B, then by the definition of a distinguished subset, C would determine both in A and in B the element $\varphi(M \setminus C)$, and then A and B would have a larger common segment than C, consisting of C and the element $\varphi(M \setminus C)$. This contradiction to the definition of C shows that one of the two distinguished subsets A and B is a segment of the other.

Hence it follows that the union L of all distinguished subsets of M is itself distinguished. For if a and b in L belong to the distinguished subsets A and B, respectively, then they both lie in the larger of the two subsets—say, in A. Upon setting $a \geqq b$ in L if $a \geqq b$ in this A, we obtain a linear ordering in L which is even a well-ordering: every descending chain of elements of L is entirely contained in some distinguished subset A and must therefore break off. Finally, if $a \in L$, then a is contained in some distinguished subset A and determines, in L and in A, one and the same segment A', where $a = \varphi(M \setminus A')$. This shows that L is distinguished.

To conclude the proof of the theorem it remains to point out that if L were different from M, then in contradiction to the definition of L we could construct a distinguished subset larger than L, by adjoining the element $\varphi(M \setminus L)$ to L and regarding this element as following all the elements of L. //

Hausdorff's Theorem follows from Zermelo's Theorem.

Let A be an arbitrary chain in a partially ordered set M. If $A = M$, then there is nothing to prove; in the contrary case, we may assume on the basis of Zermelo's Theorem that the set $B = M \setminus A$ is well-ordered; this well-ordering has nothing to do with the partial ordering of B as a subset of M.

We assign the first element of B to the *first class* if it is comparable (see § 4.1) in M with every element of A and to the *second class* otherwise. Suppose now that b is an arbitrary element of B and that every element of B strictly preceding b in the sense of the well-ordering given in B has already been assigned to the first or the second class. Then we assign b to the first class if it is comparable in M with every element of A as well as with each of those elements preceding it in B that have been assigned to the first class; otherwise, we assign b to the second class.

In this way we carry out an inductive construction in the well-ordered set B and we may therefore assume (see § 5.3) that every element of B is uniquely assigned to the first or the second class. The set C consisting of all the elements of the chain A and all the elements of the first class of B is a chain in M, because any two elements of C are comparable in M. This chain is maximal in M, because every element of the second class of B is not comparable with at least one element of C. //

The Kuratowski-Zorn Theorem follows from Hausdorff's Theorem.

For suppose that M is a partially ordered set in which every chain has an upper bound, and let $a \in M$. The chain consisting of the single element a is contained, by Hausdorff's Theorem, in a maximal chain C. If c is an upper bound of C, then $a \leqq c$. On the other hand, the element c is maximal in M: if there exists an element b for which $c < b$, then for all $x \in C$ we have $x < b$ because of $x \leqq c$; that is, by adjoining the element b to the chain C we obtain a larger chain, in contradiction to the maximality of C. //

The Axiom of Choice follows from the Kuratowski-Zorn Theorem.

Let M be an arbitrary set. We examine those systems of non-empty subsets of M on which it is possible (in at least one way) to define a function that selects an element $\varphi(A)$ in every subset A of the given system. Systems of this kind exist: for example, the systems consisting of a single non-empty set. Let Φ denote the set of all functions of the requisite kind that can be given on all possible systems of subsets on which the definition of such functions is possible.

Let φ and ψ be two functions belonging to Φ and given on the systems of subsets S and T, respectively. We set $\varphi \leqq \psi$ if $S \subseteq T$ and if the functions φ and ψ coincide on S. This defines a partial order in Φ. We now choose in Φ an arbitrary chain Γ (in the sense of this partial order) con-

sisting of functions φ_α given, respectively, on systems S_α. On the system $T=\bigcup_\alpha S_\alpha$, we can define a function ψ that coincides with φ_α on every system S_α. Clearly, ψ belongs to Φ and is an upper bound for the chain Γ.

The Kuratowski-Zorn Theorem is therefore applicable to Φ, so that Φ has maximal elements. Let χ be one of these maximal elements. If the system U on which χ is defined did not contain the non-empty subset A of M, then a function strictly larger than χ could have been defined on the system obtained from U by adjunction of A: any function will do that coincides with χ on U and selects in A one of the elements of A. This contradiction to the maximality of χ shows that U does coincide with the system of all non-empty subsets of M. //

❖ The statement that every set can be linearly ordered is weaker than the Axiom of Choice [A. Mostowski, *Fund. Math.*, vol. 32 (1932), pp. 201-52]. ❖

CHAPTER II

GROUPS AND RINGS

§ 7. Groupoids, Semigroups, and Groups

1. Basic to all the concepts that are studied in the various branches of algebra is the concept of an *algebraic operation*. For the time being we shall confine the discussion to *binary* operations. In its widest sense, a binary operation is a rule that assigns *one or more* elements of M to *certain* ordered pairs of elements of a given set M (i.e., to certain elements of the set $M \times M$; see § 2.1). If we call this operation *multiplication* and adopt the usual multiplicative notation for it, then the equation

$$ab = c \tag{1}$$

has the meaning that the product is defined for the pair of elements a, b of M and that one of its values is c.

The concept of a binary algebraic operation in this wide sense is equivalent to the concept of a ternary relation given in M (see § 2.7).

For if a binary operation is given on M, then we can introduce a ternary relation R in M by setting $R(a, b, c)$ if and only if equation (1) holds. Conversely, if a ternary relation R is given in M, then we can assume that (1) holds if and only if $R(a, b, c)$ is valid. //

2. In what follows, a *binary algebraic operation* will be understood, as a rule, in a narrower sense: the product must be defined for *every* ordered pair of elements and must be *unique*. Every set in which an algebraic operation of this type is given is called a *groupoid*.

This is still rather too wide a concept. The concept of a *semigroup*, which has diverse applications, is somewhat narrower; this is a groupoid in which the *associative law* holds: for arbitrary elements a, b, and c

$$(ab)c = a(bc). \tag{2}$$

The equation (2) gives a unique meaning to the product abc of any three elements of a semigroup. From this it follows easily that for all n the

product $a_1a_2 \ldots a_n$ of n arbitrary elements taken in a definite order is also a uniquely defined element of the semigroup.

3. Still narrower is the concept of a group, one of the most important of algebraic concepts. A *group* is a semigroup in which inverse operations can be carried out, that is, in which for arbitrary elements a and b each of the equations

$$ax = b \qquad \text{and} \qquad ya = b \tag{3}$$

has a unique solution.

We note that the uniqueness of the solution of each of the equations (3) enables us to introduce in a group a *left-hand* and a *right-hand cancellation*: if

$$ab_1 = ab_2 \qquad \text{or} \qquad b_1a = b_2a,$$

then $b_1 = b_2$.

The solutions x and y of (3) need not coincide in an arbitrary group. The fact is that we do not assume algebraic operations to be commutative, so that the product may depend on the order of the factors. A group (or a semigroup, or a groupoid) in which any two elements a, b satisfy the *commutative law*

$$ab = ba$$

is called *commutative*, or *abelian*.

4. *Every group G has a uniquely determined element e satisfying the condition*

$$ae = ea = a$$

for all elements $a \in G$.

For from the definition of a group it follows that, for every element a, G contains an element e'_a such that $ae'_a = a$ and that this element e'_a is uniquely determined. If b is any other element of G and y the element of the group that satisfies the equation $ya = b$, then by the associative law,

$$be'_a = (ya)\,e'_a = y(ae'_a) = ya = b,$$

and hence $e'_b = e'_a$. The element e'_a therefore does not depend on a; we denote it by e'. Thus,

$$ae' = a \quad \text{for all} \quad a \in G. \tag{4}$$

Similarly, we can prove the existence and uniqueness of an element e'' such that

$$e''a = a \quad \text{for all} \quad a \in G. \tag{5}$$

However, when we apply (4) and (5) to the product $e''e'$, we see that $e''e' = e''$ and $e''e' = e'$; hence $e'' = e'$. //

The element e whose existence and uniqueness is asserted in this theorem is called the *unit element* of G and is usually denoted by the symbol 1.

Every element a of G has a uniquely determined element a^{-1} such that

$$aa^{-1}=a^{-1}a=1.$$

For from the definition of a group it follows that there exist uniquely determined elements a' and a'' for which

$$aa'=1, \quad a''a=1.$$

However, by employing the associative law, we obtain

$$a''aa'=a''(aa')=a''\cdot 1=a'',$$
$$a''aa'=(a''a)a'=1\cdot a'=a',$$

and hence $a''=a'$. //

The element a^{-1} is called the *inverse element* of a. It is clear that the inverse of a^{-1} is a and that $1^{-1}=1$. It is also easy to see that for arbitrary elements $a_1, a_2, \ldots, a_n$

$$(a_1a_2\ldots a_{n-1}a_n)^{-1}=a_n^{-1}a_{n-1}^{-1}\ldots a_2^{-1}a_1^{-1}.$$

5. The following theorem often simplifies the task of verifying that a given semigroup is a group:

A semigroup G is a group if and only if G contains at least one right unit element e with the property

$$ae=a \quad \textit{for all} \quad a\in G$$

and if this e can be chosen so that every $a\in G$ has at least one right inverse element a^{-1} satisfying the condition

$$aa^{-1}=e.$$

In one direction this theorem has been proved in the preceding subsection. Now let G be a semigroup satisfying the conditions of the theorem. We shall show that the element e is also a left unit element for G. If $a\in G$ and if a^{-1} is one of its right inverse elements, then

$$eaa^{-1}=ee=e=aa^{-1}.$$

Multiplying both sides of this equation on the right by one of the elements that are right inverses of a^{-1} and making use of the uniqueness of the product in a semigroup, we obtain

$$eae=ae;$$

hence $ea=a$, as we wished to prove.

Now if e' is an arbitrary right unit element of G, and e'' an arbitrary left unit element, then, as in the proof of the first theorem of the preceding subsection, we see that $e''=e'$, i.e., we prove the existence and uniqueness of the unit element e in G.

Next, suppose again that $a\in G$ and that a^{-1} is one of the right inverse elements of a. Multiplying the equation $aa^{-1}=e$ on the left by a^{-1}, we obtain

$$a^{-1}aa^{-1}=a^{-1}.$$

Multiplying this last equation on the right by one of the right inverse elements of a^{-1}, we arrive at the equation

$$a^{-1}ae=e,$$

and hence $a^{-1}a=e$. Thus a^{-1} turns out to be a left inverse of a.

Now it is easy to verify, as in the proof of the second theorem of the preceding subsection, that every left inverse element of a is equal to every right inverse. Hence it follows that G contains, for every element a, a uniquely determined inverse element a^{-1}.

To conclude the proof of the theorem, we show that the elements

$$x=a^{-1}b \quad \text{and} \quad y=ba^{-1},$$

respectively, are solutions of the equations (3). The uniqueness of these solutions follows from the fact that if $ax_1=ax_2$, say, then by multiplying this equation on the left by a^{-1} we see that $x_1=x_2$. //

6. Occasionally, especially in the investigation of abelian groups, we use not multiplicative, but *additive* notation: the group operation is called *addition*; the sum is denoted by $a+b$; the unit element of the group is called *zero* and is denoted by the symbol 0; and instead of speaking of the inverse element, we speak of the *opposite element* and denote it by $-a$.

In the additive notation for abelian groups the inverse operation—in this case it is, of course, unique—is called *subtraction*. The solution of the equation

$$a+x=b$$

is called the *difference* of a and b and is written in the form $b-a$. Clearly,

$$b-a=b+(-a)$$

and therefore

$$b-(a_1+a_2)=b-a_1-a_2.$$

7. The usual operations on numbers give us several important examples of abelian groups. For example, by taking all the integers—positive, zero, and negative—and by considering the operation of addition in this set we obtain an abelian group, which is called the *additive group of integers*. All the rational numbers also constitute an abelian group under addition: this is the *additive group of rational numbers*. We may also speak of the additive groups of all real numbers and of all complex numbers.

Turning to the natural numbers, we see that they form a semigroup under addition—called the *additive semigroup of natural numbers*—but not a group, because subtraction is not always possible here.

In forming groups under multiplication from numbers we must beware of including the number zero, because division by zero is impossible. All the non-zero rational numbers constitute one example of a multiplicative group; the strictly positive rational numbers are another. The systems of all integers, of all non-negative integers, and of all natural numbers are semigroups, but not groups under multiplication.

As an example of a finite abelian group we mention the *multiplicative group of n-th roots of unity*. The order of this group is n, the *order* of a finite group being the number of elements in the group.

8. We now come to examples of non-commutative groups and semigroups. A *transformation* of a set M is an arbitrary mapping of the set into itself (i.e., onto one of its subsets). A special case of a transformation is a *permutation*, that is, a one-to-one mapping of M onto itself.

In § 1.2 we introduced a multiplication of mappings, interpreted as the act of carrying them out in succession, and we proved that this operation is associative. Applying this to the case of transformations, we see that under this multiplication all the transformations of a given set M form a semigroup; it is called the *symmetric semigroup on M*.

Since the result of carrying out two permutations upon a set M is itself a permutation, we may also speak of the semigroup of permutations on M. The *identity permutation*, which leaves every element of M in place, is the unit element of this semigroup. On the other hand, if x is an arbitrary permutation carrying any element a of M into ax, then the inverse transformation, carrying ax into a for all $a \in M$, is also a permutation; it is the *inverse permutation* of x. Thus, by § 7.5, all the permutations of a given set M form a group under the operation of multiplication; it is called the *symmetric group on M*.

If M is a finite set consisting of n elements, the symmetric group on M, which is called *the symmetric group of degree n* and is denoted by S_n, is a finite group of order $n!$. The symmetric semigroup on a finite set M is also finite.

Another example of a non-commutative group is the set of all non-singular square matrices of order n (where $n \geqq 2$) with real elements, when considered with respect to the operation of multiplication of matrices.

§ 8. Rings and Fields

1. Another very important algebraic concept, apart from that of a group, is the concept of a ring. A *ring* is a set R in which two binary algebraic operations (in the sense of § 7.2) are given: addition and multiplication; under the operation of addition, it must be an abelian group—the *additive group of* R—and multiplication must be connected with addition by the *distributive laws*:

$$a(b+c) = ab + ac, \qquad (b+c)a = ba + ca. \tag{1}$$

The multiplication is, in general, not subject to any restrictions; that is, under the operation of multiplication the ring R is only a groupoid—the *multiplicative groupoid of* R. If the multiplication in the ring is associative, then we speak of an *associative ring* and of its *multiplicative semigroup*; and if the multiplication in the ring is commutative as well as associative, then the ring is called *associative-commutative*.

In every ring the distributive laws also hold for subtraction, i.e.,

$$a(b-c) = ab - ac, \qquad (b-c)a = ba - ca. \tag{2}$$

For by § 7.6,

$$c + (b-c) = b.$$

Multiplying both sides of this equation on the left by a and then applying the first of the distributive laws (1) to the left-hand side we obtain

$$ac + a(b-c) = ab,$$

from which, again by § 7.6, the first of the equations (2) follows. //

2. *Every abelian group G is the additive group of a ring*: it is sufficient to write the group operation of G additively and then to introduce the *zero multiplication* into G, that is, to set

$$ab = 0$$

for arbitrary a and b of G. That the distributive laws (1) hold is obvious. This *zero ring* with additive group G is, of course, associative-commutative.

In the theory of rings the zero rings play a role parallel to that of abelian groups in the theory of groups.

As our first example of a non-zero associative-commutative ring we take the *ring of integers*. As an example of an associative, but not commutative ring we mention the *ring of square matrices* of order n (where $n \geqq 2$) with real elements, under the usual operations of addition and multiplication of matrices.

Finally, we give one example of a non-associative ring. This is the *ring of vectors of three-dimensional euclidean space* under the usual operations of addition and of vector multiplication of vectors. It is easy to verify that this multiplication is neither associative nor commutative, but that it is connected with addition by the distributive laws (1).

The reader can easily verify (or can find it proved in textbooks on vector algebra) that in our ring any three vectors a, b, c satisfy the following equation:

$$a^2 = 0 \tag{3}$$

and the *Jacobi identity*

$$(ab)c + (bc)a + (ca)b = 0. \tag{4}$$

We note that (3) implies the *anticommutative law*

$$ba = -ab.$$

For since

$$a^2 = b^2 = (a+b)^2 = 0,$$

we have

$$0 = (a+b)^2 = a^2 + ab + ba + b^2 = ab + ba.$$

3. Every ring satisfying the conditions (3) and (4) is called a *Lie ring*. Lie rings, in general, are non-associative and form an important class of rings containing, in particular, all the zero rings.

There is a noteworthy connection between associative rings and Lie rings:

If R is an arbitrary associative ring, then by preserving the additive group of the ring and replacing the operation of multiplication ab by the operation of commutation

$$a \circ b = ab - ba,$$

we obtain a Lie ring $R^{(-)}$.

For let us verify the distributive laws—for example, the first law (1):

$$a \circ (b+c) = a(b+c) - (b+c)a = ab + ac - ba - ca =$$
$$= (ab - ba) + (ac - ca) = a \circ b + a \circ c.$$

Thus, the set R with the operations of addition and commutation becomes a ring; we denote it by $R^{(-)}$. It remains to verify the equations (3) and (4):

$$a \circ a = aa - aa = 0,$$

$$\begin{aligned}(a \circ b) \circ c + (b \circ c) \circ a + (c \circ a) \circ b = \\ = (ab - ba)c - c(ab - ba) + (bc - cb)a - \\ - a(bc - cb) + (ca - ac)b - b(ca - ac) = 0.\end{aligned}$$

Thus $R^{(-)}$ turns out to be a Lie ring. //

4. *If we preserve the additive group in an associative ring R and replace the multiplication ab by the operation of symmetrization*

$$a \cdot b = ab + ba,$$

then we obtain a ring $R^{(+)}$ in which any two elements a, b satisfy the equations

$$a \cdot b = b \cdot a, \tag{5}$$

$$[(a \cdot a) \cdot b] \cdot a = (a \cdot a) \cdot (b \cdot a). \tag{6}$$

For let us verify the first distributive law (1):

$$\begin{aligned}a \cdot (b + c) = a(b + c) + (b + c)a = ab + ac + ba + ca = \\ = (ab + ba) + (ac + ca) = a \cdot b + a \cdot c.\end{aligned}$$

Next let us verify the equations (5) and (6):

$$a \cdot b = ab + ba = ba + ab = b \cdot a,$$

$$\begin{aligned}[(a \cdot a) \cdot b] \cdot a = [(aa + aa)b + b(aa + aa)]a + \\ + a[(aa + aa)b + b(aa + aa)] = aaba + \\ + aaba + baaa + baaa + aaab + aaab + \\ + abaa + abaa = (aa + aa)(ba + ab) + \\ + (ba + ab)(aa + aa) = (a \cdot a) \cdot (b \cdot a). //\end{aligned}$$

Every ring satisfying the conditions (5) and (6) is called a *Jordan ring*. Jordan rings are, in general, non-associative; they comprise, in particular, all the associative-commutative rings.

5. *In every ring R, if at least one of the factors of a product is zero, the product is itself zero. In other words,*

$$a \cdot 0 = 0 \cdot a = 0 \tag{7}$$

holds for all elements of R.

For if x is an arbitrary auxiliary element of R, then by (2),

$$a \cdot 0 = a(x - x) = ax - ax = 0. //$$

If a and b are any two elements of an arbitrary ring R, then

$$(-a)b = a(-b) = -ab, \tag{8}$$

$$(-a)(-b) = ab. \tag{9}$$

For

$$ab + (-a)b = [a + (-a)]b = 0 \cdot b = 0.$$

The second half of (8) is verified in the same way. On the other hand, using (8) we obtain

$$(-a)(-b) = -[a(-b)] = -(-ab) = ab,$$

which proves (9). //

6. We note that the converse to the statement expressed by (7) is true in the ring of integers, but is not true in general: there exist rings having *divisors of zero*, i.e., non-zero elements a and b whose product is equal to zero:

$$ab = 0.$$

Examples of such rings, apart from every zero ring, are *rings of matrices*. We have already mentioned the ring of real square matrices. Quite generally, if R is an arbitrary ring, we can consider all possible square matrices of order n with elements from R; by defining addition and multiplication in the usual way, we obtain a ring which—as is easy to verify—is even associative, provided the original ring R is associative. The zero element of this ring is the *null matrix* formed entirely of zeros. This ring is called the *full matrix ring* of order n over R and is denoted by R_n.

If $n \geqq 2$ and R does not consist of the zero element alone, then the full matrix ring R_n has divisors of zero.

For if a is a non-zero element of R, then neither of the matrices

$$\begin{Vmatrix} a & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & a \end{Vmatrix}$$

is the null matrix, but their product is. //

Other examples of rings with divisors of zero are the *full rings of functions*. Let M be an arbitrary set and R an arbitrary ring. We consider the set of all possible *functions* on M with values in R, i.e., all possible mappings f of M into R; the image of an element $x \in M$ under the mapping f is here

denoted by $f(x)$. This set of functions becomes a ring if the sum and the product of functions are defined, as usual, by the equations

$$(f+g)(x)=f(x)+g(x),$$
$$(fg)(x)=f(x)\cdot g(x),$$

that is, by addition and multiplication of the values of the given functions for all x of M. It is easy to verify that all the postulates in the definition of a ring are satisfied and that the ring so obtained is associative if the original ring R is associative, and commutative if the original ring R is commutative.

The ring just constructed is called the *full ring of functions* on M with values in R. If M is the set of points on the real line and R the set of real numbers, then our ring is the usual ring of all real functions of a real variable.

Every full ring of functions on a set M containing at least two elements, with values in a ring R that does not consist of the zero element alone, has divisors of zero.

For the role of the zero element in this ring is taken by the *zero function*—the function that is *identically zero*, that is, zero for all x of M. Now if we split M into two non-empty disjoint subsets A and B, then there obviously exist non-zero functions f and g such that f assumes the value zero on A and g the value zero on B. Clearly, the product fg is the zero function. //

7. Associative-commutative rings without divisors of zero are called integral domains; they include, in particular, all rings of numbers. Important examples of integral domains are obtained with the help of the following construction.

If R is an arbitrary associative-commutative ring, we can consider all possible *polynomials*

$$a_0+a_1x+a_2x^2+\ldots+a_nx^n \qquad (n\geqq 0)$$

in an *indeterminate* x with *coefficients* $a_0, a_1, \ldots, a_n$ from R; if $a_n\neq 0$, then n is the *degree* of this polynomial. If we define addition and multiplication of polynomials as in elementary algebra, then it is easy to verify that we obtain a ring, the so-called *ring of polynomials* $R[x]$ in the indeterminate x over R; this ring is also associative-commutative. The zero element of the ring of polynomials is the polynomial whose coefficients are all equal to zero.

Similarly, we can define the ring of polynomials $R[x_1, x_2, \ldots, x_n]$ in an arbitrary finite number of indeterminates; this is nothing but the ring of polynomials in a single indeterminate x_n over the ring $R[x_1, x_2, \ldots, x_{n-1}]$.

We can also speak of the ring of polynomials over R in arbitrary infinite set of indeterminates; here we assume that each individual polynomial involves only a finite set of indeterminates.

If R is an integral domain, then every ring of polynomials over R is also an integral domain.

That this statement is valid for the ring of polynomials $R[x]$ in a single indeterminate follows from the fact that if two polynomials f and g are different from zero and if R has no divisors of zero, then the degree of the product fg is equal to the sum of the degrees of the factors, so that the product is also different from zero. We extend this to the case of an arbitrary finite number of indeterminates by mathematical induction; and for a ring of polynomials in an infinite set of indeterminates it is sufficient to remember that every polynomial can be expressed in terms of a finite number of the indeterminates. //

8. We can examine not only polynomials over an arbitrary associative-commutative ring R, but also (formal) *power series*

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k + \ldots = \sum_{k=0}^{\infty} a_k x^k \tag{10}$$

in the indeterminate x. The definition of the operations carry over from polynomials to power series without a hitch:

$$\sum_{k=0}^{\infty} a_k x^k + \sum_{k=0}^{\infty} b_k x^k = \sum_{k=0}^{\infty} (a_k + b_k) x^k,$$

$$\sum_{k=0}^{\infty} a_k x^k \sum_{l=0}^{\infty} b_l x^l = \sum_{m=0}^{\infty} c_m x^m,$$

where

$$c_m = \sum_{k+l=m} a_k b_l.$$

As in the case of polynomials it can be verified that we obtain an associative-commutative ring; it is called the *ring of power series* in x over R and is denoted by $R\{x\}$.

The transition to rings of power series over R in an arbitrary finite number of indeterminates and then in an infinite set of indeterminates is accomplished just as for rings of polynomials; and in the case of an infinite set of indeterminates it is expedient to postulate that each power series depends on a finite number of indeterminates only.

If R is an integral domain, then every ring of power series over R is also an integral domain.

It is sufficient to discuss the case of the ring $R\{x\}$ in a single indeterminate. If we define the *lower degree* of the power series (10) as the number n for which

$$a_0 = a_1 = \ldots = a_{n-1} = 0, \qquad a_n \neq 0,$$

then every non-zero power series has a lower degree. Since there are no divisors of zero in R, the lower degree of the product of two power series is equal to the sum of the lower degrees of the factors, and our statement follows from this. //

9. The concept of a *unit element*, introduced in § 7.4 for the case of groups, carries over, of course, to the case of an arbitrary groupoid or ring R: this is an element 1 such that for every element a of R

$$a \cdot 1 = 1 \quad a = a.$$

If a ring R has one such unit element, then it has only one. There may, however, be none, as is shown by the example of the ring of even integers.

No Lie ring (see § 8.3) *with more than one element has a unit element.*

For suppose that a Lie ring L has a unit element 1. Then from (3), putting $a = 1$, we obtain $1^2 = 0$. Thus, $1 = 0$, and this implies that L consists of zero alone.

If a ring R has a unit element, then so do all possible rings of matrices over R, all rings of functions with values in R and, in the associative-commutative case, all the rings of polynomials and rings of power series over R in an arbitrary number of indeterminates.

For the unit element of a matrix ring is the unit matrix, each of whose elements along the main diagonal is 1 and whose other elements are 0. The unit element of a ring of functions is the function that is identically equal to 1. Finally, the unit elements of polynomial rings (or rings of power series) are the polynomial (power series) in which all the coefficients are equal to zero, except for a_0, which is 1. //

10. Every ring is a groupoid under multiplication, and every associative ring, a semigroup. No ring except that consisting of the zero element alone can be a group under multiplication, by virtue of the multiplicative properties (7) of zero. However, it may turn out that all the non-zero elements of a ring form a group under multiplication. Such a ring, which must necessarily be associative, is called a *field*, and the group under multiplication of its non-zero elements is the *multiplicative group* of the field. An important class of fields are the *commutative* fields. The basic examples of commutative fields are the field of rational numbers, the field of real

numbers, and the field of complex numbers. An example of a non-commutative field can be found in § 36.8.

From the definition of a field it follows that a *field does not contain divisors of zero*. Furthermore, *every field has a unit element*: for, by (7), the unit element of the multiplicative group is also that of the field. Finally, *in every field the equations*

$$ax = b, \quad ya = b, \qquad \text{where} \qquad a \neq 0, \tag{11}$$

have unique solutions.

For if $b \neq 0$, then both equations (11) have uniquely determined solutions in the multiplicative group of the field, and zero cannot satisfy either of these equations. But if $b = 0$, then zero is a solution of each of the equations (11); and there are no other solutions, because of the absence of divisors of zero. //

Conversely, *an associative ring R is a field if every equation of the form* (11) *for arbitrary $a \neq 0$ and arbitrary b has at least one solution in R.*

Let us show first of all that R has no divisors of zero. If $a \neq 0$ and $b \neq 0$, but $ab = 0$, then we let e denote one of the solutions of the equation $ax = a$ and c, one of the solutions of the equation $bx = e$. Then

$$0 = 0 \cdot c = abc = ae = a,$$

in contradiction to our assumption.

Hence it follows that the set $R^\times$ of non-zero elements of R is a multiplicative semigroup. It is even a group, because the equations (11) for $a \neq 0$ and $b \neq 0$ have solutions in $R^\times$, and the uniqueness of these solutions follows from the absence of divisors of zero in R: if, for example, $ax_1 = ax_2$ and $a \neq 0$, then $a(x_1 - x_2) = 0$ and therefore $x_1 = x_2$. //

The problem of an extension of the concept of a field to the case of non-associative rings will be treated in § 12.1.

§ 9. Subgroups and Subrings

1. A subset A of a groupoid G is called a *subgroupoid* of G if the product of any two elements of A also belongs to A. If G is a semigroup, then a subgroupoid is naturally called a *sub-semigroup*. And if G is a group, then a *subgroup* of G is a sub-semigroup A that is itself a group under the operation defined in G; for this it is sufficient that if the element a is contained in A then so is its inverse element a^{-1}.

Not every sub-semigroup of a group is necessarily a subgroup, as is shown by the example of the additive semigroup of natural numbers, which is a sub-semigroup, but not a subgroup, of the additive group of integers.

Note that it makes sense to speak of subgroups of an arbitrary semigroup. For example, the symmetric group on an arbitrary set M (see § 7.8) is a subgroup of the symmetric semigroup on that same set.

2. A *subring* of a ring R is a subset A of the ring that is itself a ring under the operations defined in R. In other words, A must be a subgroup of the additive group of R and a subgroupoid of the multiplicative groupoid of this ring: the distributive laws, being valid in R, also hold, of course, in A.

In particular, we can speak of a subring of a field. For example, the ring of integers is a subring of the field of rational numbers.

The concept of a subfield is defined similarly. We can speak of a subfield not only of a field but of an arbitrary ring. For example, in the ring of polynomials $P[x]$ over an arbitrary field P the polynomials of degree zero together with the zero element form a subfield.

3. The subgroups of an arbitrary group G include, in particular, the group itself and the *unit subgroup* E, consisting of the unit element only. Similarly, the subrings of every ring R include the ring itself and the *zero subring*, consisting of the zero element only.

In order to obtain more interesting examples of subgroups we introduce the concept of powers of an element. Let G be an arbitrary semigroup and a one of its elements. The associativity of multiplication enables us to define in the usual way the *positive powers* a^n of a ($n = 1, 2, \ldots$), where, as usual,

$$a^k \cdot a^l = a^{k+l}, \tag{1}$$

$$(a^k)^l = a^{kl}. \tag{2}$$

It follows from (1) that all the positive powers of an element form an abelian sub-semigroup of G, which is called the *cyclic sub-semigroup* of a.

If G is a group, we put $a^0 = 1$ and can then introduce *negative powers* of a. For if n is an arbitrary natural number, then it is easy to verify that

$$a^n(a^{-1})^n = 1,$$

from which it follows that

$$(a^{-1})^n = (a^n)^{-1}. \tag{3}$$

The element on both sides of (3) will be denoted by a^{-n}. The equations (1) and (2) remain valid for arbitrary powers of an element a of G, so that all the powers of a, including the zeroth and negative powers, form an abelian subgroup of G, which we call the *cyclic subgroup* of a and denote by $\{a\}$.

In the additive notation for the group operation we speak of the *multiples* of an element, instead of its powers.

4. Powers of an element a of G in which the exponents are distinct are not necessarily distinct elements of the group; the very existence of finite groups makes this clear. If all the powers of an element a are in fact distinct, then a is called an *element of infinite order,* and otherwise, an *element of finite order.* In the latter case, there exist integers k and l such that $k > l$ but

$$a^k = a^l.$$

Hence $a^{k-l} = 1$ with $k - l > 0$; in other words, there exist powers of a with positive exponents which are equal to the unit element. The smallest of these exponents is called the *order* of a.

If a is an element of finite order n, then the powers

$$1 = a^0, a, a^2, \ldots, a^{n-1} \tag{4}$$

are obviously distinct elements of the group. Every other power a^k of a, with positive or negative exponents, is equal to one of the elements (4). For if

$$k = nq + r, \qquad 0 \leqq r < n,$$

then, by (1), (2) and the equation $a^n = 1$, we have

$$a^k = (a^n)^q a^r = a^r.$$

Hence it follows that *the order n of an element of finite order a coincides with the order* (see § 7.7) *of its cyclic subgroup* $\{a\}$.

A group in which the orders of all the elements are finite, but not necessarily bounded, is called *periodic.* At the other extreme, a group is called *torsion-free* if all its elements except 1 are of infinite order.

5. In the case of rings, the problem naturally arises of an analogue of a cyclic subgroup, that is, of the minimal subring containing a given element. Let R be an arbitrary (not necessarily associative) ring and a one of its elements. Every subring of R containing a also contains all possible products of n factors each equal to a ($n = 1, 2, 3, \ldots$); these products play the role of the positive powers of a. Since the multiplication is possibly nonassociative, in every such product parentheses must be inserted in such a way that only two elements of the ring are multiplied each time. For $n = 3$, there are two such products, $(aa)a$ and $a(aa)$; for $n = 4$, there are five; and so forth.

Every subring containing a also contains all possible sums of an arbitrary finite number of such products taken with arbitrary integer coefficients. However, the elements of R that can be written in the form of such sums

(possibly not in a unique way) themselves form a subring of R; this, then, is the required *subring generated by a.*

From what we have said, it follows that if R is an associative ring, then the subring generated by a consists of all those elements of R that can be written in at least one way as sums (with integer coefficients) of positive powers of a. This subring is therefore commutative.

6. Every subgroup of a group G contains the unit element of the group; every subring of a ring R, the zero element of the ring. The intersection of an arbitrary system of subgroups of G (or subrings of R) is therefore not empty.

The intersection of an arbitrary system of subgroups of a group G is a subgroup. Similarly, the intersection of an arbitrary system of subrings of a ring R is a subring, the intersection of an arbitrary system of sub-semigroups of a semigroup, provided it is not empty, is a sub-semigroup, etc.

We shall only prove the first statement. In a group G let us take an arbitrary system of subgroups A_α, where α ranges over an index set, and let D be the intersection of these subgroups. If b and c are any two elements of D, then their product bc is contained in each of the subgroups A_α, and hence in D. Furthermore, for every $b \in D$ the element b^{-1} belongs to each of the subgroups A_α, so that $b^{-1} \in D$. The set D is consequently a subgroup of G. //

7. Let G be a group or a ring or a field or a semigroup or an algebraic system of any other type we have met with thus far. If M is an arbitrary subset of G, then there exists a minimal subgroup (subring, subfield, etc.) containing M, the *subgroup (subring, subfield,* etc.) *generated by M*; it is denoted by $\{M\}$. This is precisely the intersection of all subgroups (subrings, subfields, etc.) containing the whole of M; at least one such exists, namely G itself.

By generalizing the definition of cyclic subgroups given in § 9.3 we can easily verify that *the subgroup $\{M\}$ consists of those and only those elements of G that can be written in at least one way in the form of a product of powers of a finite number of elements of M.*

Similarly, generalizing the results of § 9.5, we see that *the subring generated by a set M in a ring R consists of all those elements of R that can be written in at least one way in the form of a sum (with integer coefficients) of an arbitrary finite number of elements of M*; in the non-associative case these products are, of course, provided with a certain distribution of parentheses.

8. By applying the remarks above to the case where a certain system of subgroups A_i $(i \in I)$ is given in a group G and the set M is the union of all

the A_i, we arrive at the concept of the *subgroup generated by a given system of subgroups*; this is denoted by $\{A_i, i \in I\}$ in the general case, by $\{A, B\}$ when we are dealing with the union of two subgroups A and B, etc. The concept of *the subring generated in a ring by a given system of subrings* has a similar meaning.

9. If G is a group or a ring, etc., then there exist subsets M—for example, G itself—such that

$$\{M\} = G.$$

Every such subset M is called a *system of generators* for G.

If G has at least one finite system of generators, then we say that G is a *finitely generated group* (semigroup, ring, etc.). A system of generators may consist, in particular, of a single element. Groups that have one generating element, i.e., that coincide with one of their cyclic subgroups, are called *cyclic groups*. A complete survey of cyclic groups will be given in the next section.

§ 10. Isomorphism

1. The concept of isomorphism can be introduced for each of the types of algebraic systems discussed in this chapter, and it plays the same role here as for partially ordered sets (see § 4.3): for example, isomorphic groups may be regarded as identical—as two copies of one and the same group—whenever the group operation is the primary object of study and the nature of the elements forming the groups is unimportant.

We now come to the definitions. Two groupoids G and G' are called *isomorphic* if there exists a one-to-one mapping φ of G onto G' such that for arbitrary elements $a, b \in G$

$$(ab)\varphi = a\varphi \cdot b\varphi.$$

A mapping φ with these properties is called an *isomorphism*. Clearly, the property of groupoids of being isomorphic is symmetric (the mapping inverse to an isomorphism is itself an isomorphism) and transitive; it is also reflexive: it suffices to consider the identity mapping of a groupoid onto itself. An isomorphism of the groupoids G and G' is usually expressed by the symbol

$$G \cong G'.$$

We shall use the same symbol to express an isomorphism in the case of other algebraic structures.

It is clear that an isomorphic mapping preserves all the properties of the groupoid that can be stated in the language of the operation given in the

groupoid, such as associativity, commutativity, and the existence of a unit element and inverse elements. Let us show by a simple example how statements of this kind can be proved. Let G be a commutative groupoid and φ an isomorphic mapping of G onto a groupoid G'. If a' and b' are any two elements of G' and a and b are elements of G for which

$$a\varphi = a', \quad b\varphi = b',$$

then

$$(ab)\varphi = a'b', \quad (ba)\varphi = b'a',$$

and the equation $ab = ba$ together with a uniqueness of the mapping imply that $a'b' = b'a'$.

Hence it follows that *an isomorphic image of a semigroup, group, and abelian group is itself a semigroup, group, and abelian group, respectively*

Two rings are called *isomorphic* if a one-to-one correspondence can be set up between them that is an isomorphism of the additive groups as well as of the multiplicative groupoids of the rings. Clearly, an isomorphism of rings preserves the property of being associative-commutative, of being a Lie ring or a Jordan ring, and of being a field.

Let us point out one interesting example of an isomorphism of groups: *The multiplicative group of positive real numbers is isomorphic to the additive group of all real numbers.* For when we assign to every positive number its logarithm to a fixed base, we obtain a one-to-one mapping of the first of these groups onto the second. That it is an isomorphism follows from the fact that the logarithm of a product is equal to the sum of the logarithms of the factors. //

2. We now proceed to the survey of all cyclic groups promised in § 9.9 and we begin with some examples. The additive group of integers is an infinite cyclic group, because every integer is a multiple of 1. Next, the multiplicative group of the n-th roots of unity is a finite cyclic group of order n, a fact that follows from the existence of primitive n-th roots of unity.

All infinite cyclic groups are isomorphic to the additive group of integers and are therefore isomorphic to each other. All finite cyclic groups of order n are isomorphic to the multiplicative group of the n-th roots of unity and are therefore isomorphic to each other.

Let us prove the first of these statements. If G is an infinite cyclic group with a generating element a, then the correspondence

$$a^k \to k$$

is a one-to-one mapping of G onto the whole additive group of integers. That this is an isomorphism follows from § 9.3, (1).

To prove the second statement it is sufficient to set up the correspondence between powers (with equal exponents) of a generating element of the given cyclic group and an appropriate primitive root of unity. //

In virtue of this theorem, for the purpose of obtaining, say, a survey of all subgroups of an infinite cyclic group, it is sufficient to examine the additive group of integers.

Every non-zero subgroup of the additive group of integers consists entirely of the multiples of a certain natural number n.

For it is clear that all the integers that are multiples of the natural number n form a subgroup, namely the cyclic subgroup of n, and that for distinct n these subgroups are distinct. On the other hand, if A is an arbitrary non-zero subgroup of the additive group of integers, then it does not consist of negative numbers only, because for every number it contains it also contains its opposite. Let n be the smallest natural number contained in A. If a is an arbitrary number from A, let

$$a = qn + r, \qquad 0 \leqq r < n.$$

Then

$$r = a - qn \in A,$$

and therefore, on account of the way n was chosen, we have $r = 0$, so that $a = qn$. //

Thus, *every subgroup of an infinite cyclic group is itself cyclic.* The same is true for finite cyclic groups.

3. We say that a groupoid G is *isomorphically embedded* in a groupoid G' if there exists an isomorphic mapping of G onto a subgroupoid of G'. This concept carries over, of course, to the case of rings.

Every ring R can be isomorphically embedded in the full matrix ring R_n (see § 8.6).

For the *scalar matrices*—that is, the matrices that have one and the same element a along the main diagonal and zeros everywhere else—form a subring R_n that is isomorphic to R. //

Every ring R can be isomorphically embedded in the full ring of functions over a given set M with values in R (see § 8.6).

For the functions that assume one and the same value $a \in R$ for all x of M form a subring isomorphic to R in the ring of functions. //

For an arbitrary associative-commutative ring R, the ring of polynomials $R[x]$ can be isomorphically embedded in the ring of power series $R\{x\}$ (see § 8.8).

For, the power series that have only a finite number of non-zero coefficients form a subring of $R\{x\}$ isomorphic to $R[x]$. //

4. A slightly greater effort is required for the proof of the following theorem:

Every ring R can be isomorphically embedded in a ring with unit element. If R is associative or commutative, then it can be embedded in an associative or commutative ring, respectively, with unit element.

Proof: We examine the set of all possible pairs of the form (a, k), where $a \in R$ and k is an integer. We define the sum and product of such pairs by the equations

$$(a, k) + (b, l) = (a + b, k + l), \tag{1}$$

$$(a, k)(b, l) = (ab + la + kb, kl). \tag{2}$$

Under addition, we obviously obtain an abelian group. Since

$$\begin{aligned}[(a, k) + (b, l)](c, m) &= (a + b, k + l)(c, m) = \\ &= ((a + b)c + m(a + b) + (k + l)c,\ (k + l)m) = \\ &= (ac + ma + kc + bc + mb + lc,\ km + lm) = \\ &= (ac + ma + kc, km) + (bc + mb + lc, lm) = \\ &= (a, k)(c, m) + (b, l)(c, m),\end{aligned}$$

and since the second distributive law can be verified in the same way, we have constructed a ring. From (2) it follows that

$$(a, k)(0, 1) = (0, 1)(a, k) = (a, k),$$

i.e., the pair $(0, 1)$ is the unit element of the ring. Finally, by (1) and (2),

$$(a, 0) + (b, 0) = (a + b, 0),$$

$$(a, 0)(b, 0) = (ab, 0),$$

i.e., in our ring of pairs, the pairs of the form $(a, 0)$ form a subring isomorphic to R.

This proves the first statement of the theorem. The second statement follows easily from (2). //

Note that the ring we have constructed is by no means the unique (or minimal) ring with the unit element containing (in the sense of isomorphic embedding) the given ring R: in fact, R itself may have a unit element.

5. *Every groupoid can be isomorphically embedded in the multiplicative groupoid of a ring; and an associative or a commutative groupoid can be embedded in a ring having that same property.*

To prove this, we consider all possible sums of the form

$$\sum_{a \in G} k_a a, \tag{3}$$

where a ranges over all the elements of the given groupoid G and the coefficients k_a are integers, and where *only a finite number of these coefficients is different from zero.* We define addition and multiplication of sums of the form (3) in the following way:

$$\sum_{a \in G} k_a a + \sum_{a \in G} l_a a = \sum_{a \in G} (k_a + l_a) a, \tag{4}$$

$$\sum_{a \in G} k_a a \cdot \sum_{b \in G} l_b b = \sum_{c \in G} m_c c, \tag{5}$$

where m_c is the sum of all non-zero products $k_a l_b$ over those a and b for which $ab = c$. It is clear that the right-hand sides of (4) and (5) are themselves sums of the form (3); the equation (5) has the following meaning: the finite sums occurring as factors on the left-hand side are to be multiplied termwise, with

$$k_a a \cdot l_b b = (k_a l_b)(ab),$$

where the product ab is to be understood in the sense of the operation given in G, and like terms are then to be collected.

Obviously, we obtain an abelian group under addition. The verification of the distributive laws and the proofs that associativity or commutativity of the multiplication in G implies the same property for the multiplication of sums of the form (3) are a little tedious, but do not present any fundamental difficulties, and we leave them to the reader.

Thus, all the sums (3) form a ring under the operations defined by (4) and (5). It follows from (5) that the sums of the form (3) in which $k_a = 1$ for exactly one element a of G and all the remaining coefficients are zero form a subgroupoid of the multiplicative groupoid of this ring that is isomorphic to the given groupoid G. //

The ring constructed above is called the *integral groupoid ring* of G. If G is a semigroup or a group, then we speak of the *integral semigroup ring* and *the integral group ring*, respectively.

6. *Every group G can be isomorphically embedded in the symmetric group on a certain set M* (§ 7.8). *M can be taken to be the set of elements of G itself.*

For let us put into correspondence with every element a of G the transformation of this group that carries an arbitrary element x of G into xa.

Since $x \neq y$ implies that $xa \neq ya$ and since for every x of G the equation

$$(xa^{-1})a = x$$

holds, the transformation $x \to xa$ is a one-to-one mapping of G onto itself, that is, a permutation. Further, if $a \neq b$, then the permutations corresponding to these elements are also distinct, because (for example) $1 \cdot a \neq 1 \cdot b$. Finally, the equation

$$(xa)b = x(ab)$$

shows that the permutation corresponding to the product ab coincides with the result of carrying out in succession the permutations corresponding to a and to b. //

Every semigroup G can be isomorphically embedded in the symmetric semigroup on a certain set M (see § 7.8).

To prove this, we use § 10.5 to embed G isomorphically in the multiplicative semigroup of an associative ring R; we then embed this ring, by § 10.4, in an associative ring $\overline{R}$ with a unit element and we denote the multiplicative semigroup of this latter ring by $\overline{G}$. Since $\overline{G}$ contains an element 1, for which $a \neq b$ implies that $1 \cdot a \neq 1 \cdot b$, we can repeat the proof of the preceding theorem almost verbatim and we see that the semigroup $\overline{G}$ can be isomorphically embedded in the symmetric semigroup on the set $\overline{G}$ itself. This then yields the required embedding for the given semigroup G. //

Observe that the possibility of embedding an arbitrary groupoid (in particular, an arbitrary semigroup) in a groupoid (semigroup) with a unit element could also be proved directly, without recourse to rings.

§ 11. Embedding of Semigroups in Groups and of Rings in Fields

1. Not every semigroup G can be isomorphically embedded in a group: a necessary condition for this is that the *cancellation law* holds in G (see § 7.3):

From $ac = bc$, and also from $ca = cb$, it follows that $a = b$.

It is very easy to construct an example of a semigroup that violates this condition.

Similarly, not every associative ring R can be isomorphically embedded in a field: for this it is necessary (see § 8.10) that R has no divisors of zero.

These necessary conditions are, in general, by no means sufficient.

❖ Necessary and sufficient conditions for a semigroup to be embeddable in a group can be expressed in the form of an infinite set of postulates of

the form: "a given system of equations implies such and such an equation"; they cannot be expressed by means of a finite number of such postulates [A. I. Mal'cev, *Mat. Sb.*, vol. 6 (1939), pp. 331-36; vol. 8 (1940), pp. 251-64].

There exist associative rings without divisors of zero that cannot be embedded in a field [A. I. Mal'cev, *Math. Ann.*, vol. 113 (1937), pp. 686-91]. ❖

2. The object of the present subsection is to prove that *the necessary conditions given above for a semigroup to be embeddable in a group, and an associative ring in a field, are also sufficient in the commutative case.* We begin by proving the following auxiliary statement:

Suppose that in an abelian semigroup G a sub-semigroup S is selected and that cancellation by elements of S is possible in G, i.e., that $ax = bx$ ($x \in S$ and $a, b \in G$) always implies that $a = b$. Then G can be isomorphically embedded in an abelian semigroup $\overline{G}$ with unit element such that every element of S has an inverse in $\overline{G}$.

We form the set of all possible *fractions of the form* $\frac{a}{x}$ ($a \in G$, $x \in S$), where this is to be interpreted simply as the ordered pair of elements (a, x). Two fractions $\frac{a}{x}$ and $\frac{b}{y}$ will be regarded as equal

$$\frac{a}{x} = \frac{b}{y}$$

if and only if $ay = bx$. This relation of equality is obviously reflexive and symmetric. It is also transitive, because if

$$\frac{b}{y} = \frac{c}{z}$$

that is, $bz = cy$, then

$$ayz = bxz = cyz,$$

from which, after cancellation of $y \in S$, we obtain $az = cx$, or

$$\frac{a}{x} = \frac{c}{z}$$

Hence (see § 3.2) the whole set of fractions splits into disjoint classes of equal fractions. We denote the set of these classes by $\overline{G}$.

We define a *multiplication of fractions* by the equation

$$\frac{a}{x} \cdot \frac{b}{y} = \frac{ab}{xy}. \tag{1}$$

This definition is meaningful, because $xy \in S$.

If

$$\frac{a}{x} = \frac{a_1}{x_1}, \quad \frac{b}{y} = \frac{b_1}{y_1}, \tag{2}$$

that is,

$$ax_1 = a_1x, \qquad by_1 = b_1y, \tag{3}$$

then

$$abx_1y_1 = a_1b_1xy,$$

and hence

$$\frac{ab}{xy} = \frac{a_1b_1}{x_1y_1}.$$

Thus, when we replace the factors on the left-hand side of (1) by fractions equal to them, we obtain a product that is equal to the right-hand side of (1).

This enables us to regard (1) as the definition of a multiplication of classes of equal fractions. The associativity and commutativity of this multiplication are obvious, and we have thus converted $\bar{G}$ into an abelian semigroup.

All the fractions of the form $\frac{z}{z}$ $(z \in S)$ are equal to each other. On the other hand, if

$$\frac{a}{x} = \frac{z}{z},$$

then $az = zx$, so that after cancellation of z, we have $a = x$. The fractions of the form $\frac{z}{z}$ therefore form a class by themselves. This class plays the role of the unit element of $\bar{G}$. Indeed, common factors belonging to S may be cancelled in a fraction:

$$\frac{az}{xz} = \frac{a}{x} \quad (z \in S);$$

hence by $(az)x = a(xz)$ we have

$$\frac{a}{x} \cdot \frac{z}{z} = \frac{az}{xz} = \frac{a}{x}.$$

If a is a fixed element of G, then all the fractions of the form $\frac{ax}{x}$ are equal to each other: from

$$(ax)y = (ay)x$$

it follows that

$$\frac{ax}{x} = \frac{b}{y}.$$

On the other hand, if

$$\frac{ax}{x} = \frac{b}{y},$$

then $axy = bx$, or $b = ay$. Finally, if

$$\frac{ax}{x} = \frac{by}{y},$$

then $axy = byx$, and hence $a = b$. Thus, by assigning to every element a of G the class of equal fractions of the form $\frac{ax}{x}$, we obtain a one-to-one mapping of G into $\bar{G}$. The fact that this is an isomorphism follows from the equation

$$\frac{ax}{x} \cdot \frac{by}{y} = \frac{(ab)(xy)}{xy}.$$

To complete the proof, we observe that the elements $z \in S$ corresponds to the class of fractions of the form $\frac{zy}{y}$. The existence of an inverse element of this class in $\bar{G}$ is a consequence of the following remark:

The class of fractions equal to $\frac{x}{y}$, where $x, y \in S$, has an inverse element in $\bar{G}$. This is the class of fractions equal to $\frac{y}{x}$.

In fact,

$$\frac{x}{y} \cdot \frac{y}{x} = \frac{xy}{yx} = \frac{t}{t} \quad (t \in S). \text{ //}$$

3. The proposition proved above and the last remark of the preceding subsection lead to the following theorem (the case $S = G$):

Every abelian semigroup with cancellation law can be isomorphically embedded in an abelian group.

By applying our construction to the additive semigroup of natural numbers, which is an abelian semigroup with cancellation law, the reader can obtain the usual embedding of this semigroup in the additive group of integers.

4. *Let N be a set of non-zero elements in an associative-commutative ring R that are not divisors of zero. Then R can be isomorphically embedded in an associative-commutative ring $\bar{R}$ with unit element such that every element of N has an inverse in $\bar{R}$.*

We observe first of all that the sub-semigroup S of the multiplicative semigroup of R generated by N (see § 9.7) also contains neither zero nor

divisors of zero. For S consists of all possible products of elements of N. However, if $a \neq 0$, $b \neq 0$ and both these elements are not divisors of zero, then $ab \neq 0$ and from $(ab)c = 0$ it follows that $bc = 0$ and hence that $c = 0$.

Thus, cancellation by elements of S is possible in the multiplicative semigroup of R and so, by § 11.2, this semigroup can be embedded in a commutative semigroup $\bar{R}$ with unit element in which every element of S and, in particular, of N has an inverse. We shall assume that the semigroup $\bar{R}$ is obtained by the construction explained in § 11.2.

As we wish to make $\bar{R}$ into a ring, we define an *addition of fractions* of the form $\frac{a}{x}$ $(a \in R,\ x \in S)$ by the equation

$$\frac{a}{x} + \frac{b}{y} = \frac{ay + bx}{xy}. \tag{4}$$

This definition is meaningful, because $xy \in S$.

If the equations (2) and hence (3) hold, then

$$(ay + bx)x_1y_1 = ayx_1y_1 + bxx_1y_1 = a_1xyy_1 + b_1yxx_1 = (a_1y_1 + b_1x_1)xy,$$

that is,

$$\frac{ay + bx}{xy} = \frac{a_1y_1 + b_1x_1}{x_1y_1}.$$

Thus, (4) can be regarded as a definition of addition of classes of equal fractions, i.e., as addition in $\bar{R}$. That this addition is commutative is obvious, and that it is associative can easily be verified on the basis of (4).

All fractions of the form $\frac{0}{z} = \frac{0z}{z}$ are equal to each other and form a complete class of equal fractions. This class plays the role of the zero element in $\bar{R}$, because

$$\frac{a}{x} + \frac{0}{z} = \frac{az}{xz} = \frac{a}{x}.$$

Every element of $\bar{R}$ has an opposite element, because

$$\frac{a}{x} + \frac{(-a)}{x} = \frac{ax + (-a)x}{x^2} = \frac{0}{x^2}.$$

Thus we see that $\bar{R}$ is an abelian group under addition.

Addition and multiplication in $\bar{R}$ are connected by the distributive law:

$$\frac{a}{x} \cdot \frac{c}{z} + \frac{b}{y} \cdot \frac{c}{z} = \frac{ac}{xz} + \frac{bc}{yz} = \frac{acyz + bcxz}{xyz^2} = \frac{acy + bcx}{xyz} =$$

$$= \frac{ay + bx}{xy} \cdot \frac{c}{z} = \left(\frac{a}{x} + \frac{b}{y}\right) \cdot \frac{c}{z}$$

Thus, $\bar{R}$ is an associative-commutative ring. As we know, the mapping that carries every element $a \in R$ into the class of equal fractions of the form $\frac{ax}{x}$ is isomorphic under multiplication. It is also isomorphic under addition, as the equation

$$\frac{ax}{x} + \frac{by}{y} = \frac{axy + byx}{xy} = \frac{(a+b)xy}{xy}$$

shows. //

The ring $\bar{R}$ constructed above is called the *ring of fractions* of R with respect to the set N of non-divisors of zero (or the multiplicative semigroup S of non-divisors of zero). As we know from § 11.2, *in the ring of fractions* $\bar{R}$, *inverse elements exist not only for all the elements of* S, *but for all the elements that can be represented as fractions of the form* $\frac{x}{y}$ $(x, y \in S)$.

5. Let us apply the results thus obtained to the case where R is an integral domain (see § 8.7). In this case we can put $N = R \setminus 0$ and $S = N$. Thus we come to the very important theorem:

Every integral domain R *can be isomorphically embedded in a commutative field.*

For in this case every non-zero element of $\bar{R}$ can be written in the form $\frac{x}{y}$ $(x, y \in S)$, so that every such element has an inverse in $\bar{R}$. The ring $\bar{R}$ is therefore a commutative field, the *field of fractions* of the integral domain R. //

❖ Consider an associative ring R without divisors of zero that satisfies the following condition (which always holds in the commutative case): For arbitrary non-zero elements a and b of R it is possible to find non-zero elements x and y in R such that $ax = by$. Then R can be isomorphically embedded in a field [O. Ore, *Ann. of Math.* vol. 32 (1931), pp. 463-77]. ❖

6. The existence theorem for the field of fractions proved above can be supplemented by the following uniqueness theorem:

Let R *be an integral domain and* $\bar{R}$ *its field of fractions; moreover, let* R *be a subring of a commutative field* P *such that the subfield of* P *generated by* R *coincides with* P *itself. Then there exists an isomorphic correspondence* φ *between* $\bar{R}$ *and* P *that maps the ring* R *identically onto itself, and this correspondence* φ *is uniquely determined.*

Let us denote an element a of R, regarded as an element of P, by a'. Then the correspondence $a \to a'$ is the identity mapping of R onto itself.

We shall prove, to begin with, that if an isomorphic mapping φ of $\bar{R}$ onto P that is the identity on R exists at all, then it is uniquely determined. For if $\frac{a}{b}$ is an arbitrary fraction, then from

$$b \cdot \frac{a}{b} = a$$

it follows that

$$b\varphi \cdot \left(\frac{a}{b}\right)\varphi = a\varphi,$$

i.e., taking $a\varphi = a'$, $b\varphi = b'$ into account,

$$b' \cdot \left(\frac{a}{b}\right)\varphi = a'.$$

Thus, the isomorphism φ carries the element $\frac{a}{b}$ of $\bar{R}$ into the quotient of the elements a' and b' of P and is therefore uniquely determined.

We now come to the proof of the main statement of the theorem. Let us assign to the element $\frac{a}{b}$ of $\bar{R}$ the quotient $\frac{a'}{b'}$ of the elements a' and b' of P, and let us set $\frac{a'}{b'} = \left(\frac{a}{b}\right)\varphi$. If $\frac{a}{b} = \frac{c}{d}$, that is, $ad = bc$, then the equation $a'd' = b'c'$ holds in P, so that from $b' \cdot \frac{a'}{b'} = a'$ we have

$$b'd' \cdot \frac{a'}{b'} = a'd' = b'c';$$

hence

$$d' \cdot \frac{a'}{b'} = c',$$

that is, $\frac{a'}{b'} = \frac{c'}{d'}$. Therefore the mapping φ does not depend on how an element of $\bar{R}$ is expressed in the form of a fraction, i.e., it is a single-valued mapping of $\bar{R}$ onto P. It is even a one-to-one mapping, because if an equation $\frac{a'}{b'} = \frac{c'}{d'}$ holds in P, then $a'd' = b'c'$, and therefore in R we have $ad = bc$, so that the equation $\frac{a}{b} = \frac{c}{d}$ holds in the field $\bar{R}$.

The isomorphism of the mapping φ follows from the fact that the equations (1) and (4), which define multiplication and addition of fractions, necessarily hold for quotients in an arbitrary commutative field P.

We have thus obtained an isomorphic mapping of the field of fractions $\bar{R}$ into P. This mapping is the identity on R, because for every $a \in R$

$$\left(\frac{ab}{b}\right)\varphi = \frac{a'b'}{b'} = a' \in P.$$

Finally, we know that the isomorphic image of a commutative field is always a field. Therefore, those elements of P that can be written in the form of fractions of elements of R form a subfield in P containing the whole ring R and therefore, by the conditions of the theorem, coinciding with P. We have thus constructed an isomorphic mapping φ of $\bar{R}$ onto the whole field P, a mapping that is the identity on R. //

7. The field of fractions of the ring of integers is the field of rational numbers.

If P is an arbitrary commutative field, then the ring of polynomials $P[x]$ is an integral domain (see § 8.7) and can therefore be embedded in a field of fractions. This field is denoted by $P(x)$ and is called the *field of rational fractions* in the indeterminate x over P. Its elements have the form of fractions $\frac{f(x)}{g(x)}$, where $f(x)$ and $g(x)$ are polynomials from $P[x]$ and $g(x) \neq 0$, and where equality of these fractions and the operations on them are defined in accordance with §§ 11.2 and 11.4.

The concept of the field of rational fractions $P(x_1, x_2, \ldots, x_n)$ in several indeterminates has an analogous meaning.

❖ The field of fractions of the ring of power series $P\{x\}$ over a commutative field P (see § 8.8) is isomorphic to the *field of 'Laurent' series* over P, that is, series of the form

$$a_n x^n + a_{n+1} x^{n+1} + \ldots, \tag{5}$$

where n may be greater than, equal to, or less than zero and all the coefficients belong to P; the series (5) therefore contains only a finite number of terms with negative powers of the indeterminate x and an infinite number of terms with positive powers. The operations on these series are performed according to rules that are natural generalizations of the rules for the operations in $P\{x\}$. ❖

§ 12. Non-Associative Fields, Quasigroups. Isotopy

1. The concept of a field can be extended to the non-associative case in several nonequivalent ways. For example, we can examine rings in which for any two elements a and b with $a \neq 0$ the equations

$$ax = b, \qquad ya = b \tag{1}$$

have solutions, but solutions that are not necessarily uniquely determined. We shall call every such ring a *division ring*; it follows from our definition that a division ring may have divisors of zero.

A division ring in which the equations (1) have *unique* solutions is called a *quasifield*. A quasifield cannot contain divisors of zero, so that the non-zero elements of a quasifield form a groupoid under multiplication. This is a *quasigroup*, that is, a groupoid in which the equations (1) are uniquely solvable for arbitrary elements a and b.

Finally, we reserve the term (*non-associative*) *field* for a quasifield having a unit element. The non-zero elements of such a field form, under multiplication, a quasigroup with unit element—a so-called *loop*.

❖ Every (not necessarily associative) ring without divisors of zero can be embedded in a quasifield [B. H. Neumann, *Proc. London Math. Soc.* vol. 1 (1951), pp. 241-56]. ❖

2. There exist quasifields that are not fields and quasigroups that are not loops. The fact of the matter is that although the equations

$$ax = a, \qquad ya = a \tag{2}$$

are uniquely solvable in a quasigroup G for every element a, the solutions do not necessarily coincide and need not be right or left unit elements for other elements of G.

Every subquasigroup A of a loop G with unit element e is a subloop with the same unit element e.

In fact, for every $a \in A$ the equations (2) have in G the unique solution e, which must therefore be contained in the subquasigroup A. //

3. A number of results concerning groups can be extended to quasigroups and, in particular, to loops. We shall not dwell further on this and shall confine ourselves to introducing a generalization of the concept of isomorphism that plays an appreciable role in the theory of quasigroups.

Let G be a groupoid with multiplication $a \cdot b$ and φ, ψ, and χ arbitrary one-to-one mappings of G onto itself, not necessarily distinct. We obtain a new groupoid on the set G by putting, for arbitrary $a, b \in G$,

$$a \circ b = (a\varphi \cdot b\psi)\chi. \tag{3}$$

This new groupoid need not be isomorphic to the old one, but they are close to a certain extent.

Accordingly, we call a groupoid G with multiplication $a \circ b$ *isotopic* to a groupoid G' with multiplication $a' \cdot b'$ if there exist three one-to-one mappings φ, ψ, and χ^{-1} of G onto G' such that for arbitrary $a, b \in G$

$$(a \circ b)\chi^{-1} = a\varphi \cdot b\psi. \tag{4}$$

Clearly, *for* $\varphi = \psi = \chi^{-1}$ *we obtain an isomorphic mapping of* G *onto* G'.

It is easy to verify that the relation of isotopy is reflexive, transitive, and symmetric.

In many cases it is convenient to assume that both the operations $a \circ b$ and $a \cdot b$ are given on one and the same set G and that φ, ψ, and χ occurring in (3) or (4) are one-to-one mappings of this set onto itself.

4. *Every groupoid that is isotopic to a quasigroup is itself a quasigroup.*

For suppose that on a set G there are given a groupoid with multiplication $a \circ b$ and a quasigroup with multiplication $a \cdot b$ and that these are isotopic, i.e., that (3) holds. Let us show that, say, the equation

$$a \circ x = b \tag{5}$$

has a unique solution for arbitrary $a, b \in G$. We know that the equation

$$a\varphi \cdot y = b\chi^{-1}$$

has a unique solution c. We set $x = c\psi^{-1}$. Then

$$a \circ x = a \circ c\psi^{-1} = (a\varphi \cdot c)\chi = (b\chi^{-1})\chi = b.$$

On the other hand, if x' is an arbitrary solution of (5), then

$$a\varphi \cdot x'\psi = b\chi^{-1},$$

and hence $x'\psi = c$, or $x' = c\psi^{-1}$. //

5. *Every quasigroup is isotopic to a loop* [A. A. Albert, *Trans. Amer. Math. Soc.*, vol. 54 (1943), pp. 507-19].

For let us fix an arbitrary element e in a quasigroup G with multiplication $a \cdot b$. Then there exists an element f in G for which

$$e \cdot f = e. \tag{6}$$

Let a be an arbitrary element of G. We denote by $a\varphi$ and $a\psi$ the uniquely determined elements for which

$$a\varphi \cdot f = a, \; e \cdot a\psi = a. \tag{7}$$

The mappings $a \to a\varphi$ and $a \to a\psi$, where a ranges over the whole set G, are one-to-one mappings of G onto itself. Thus, if $a\varphi = b\varphi$, then $a\varphi \cdot f = b\varphi \cdot f$, so that $a = b$. On the other hand, if c is an arbitrary element of G, then

$$c = (c \cdot f)\varphi.$$

Hence it follows that by setting, for all $a, b \in G$,

$$a \circ b = a\varphi \cdot b\psi, \tag{8}$$

we define a new groupoid on G that is isotopic to the original quasigroup; the role of χ is played here by the identity transformation, that is, we have what we call a *principal isotope*. By what has been proved in the preceding section, the groupoid so obtained is itself a quasigroup. It is even a loop, because e serves as a unit element for it: since by (6) and (7)

$$e\varphi = e, \qquad e\psi = f,$$

we have by (8) and (7)

$$a \circ e = a\varphi \cdot f = a,$$

$$e \circ a = e \cdot a\psi = a. \text{ //}$$

6. *If a groupoid with a unit element is isotopic to a semigroup, then they are isomorphic and therefore both are associative and both have a unit element* [R. H. Bruck, *Trans. Amer. Math. Soc.*, vol. 60 (1946), pp. 245-354; N. J. S. Hughes, *J. London Math. Soc.*, vol. 32 (1957), pp. 510-11].

For suppose that on a set G there are given a groupoid with multiplication $a \cdot b$ having a unit element e and a semigroup with multiplication $a \circ b$, and that they are isotopic:

$$a \circ b = (a\varphi \cdot b\psi)\chi,$$

where φ, ψ, χ are one-to-one mappings of G onto itself. Since for arbitrary $a, b, c \in G$

$$(a \circ b) \circ c = a \circ (b \circ c),$$

we have

$$[(a\varphi \cdot b\psi)\chi\varphi \cdot c\psi]\chi = [a\varphi \cdot (b\varphi \cdot c\psi)\chi\psi]\chi;$$

and hence

$$(a\varphi \cdot b\psi)\chi\varphi \cdot c\psi = a\varphi \cdot (b\varphi \cdot c\psi)\chi\psi. \tag{9}$$

Setting $a\varphi = c\psi = e$ in this equation, we obtain for all $b \in G$

$$b\psi\chi\varphi = b\varphi\chi\psi. \tag{10}$$

Next, by setting $a\varphi = e$ in (9) and using (10), we obtain

$$b\varphi\chi\psi \cdot c\psi = (b\varphi \cdot c\psi)\chi\psi$$

or, after replacing $b\varphi$ by a and $c\psi$ by b,

$$a\chi\psi \cdot b = (a \cdot b)\chi\psi \tag{11}$$

for all $a, b \in G$. Finally, setting $c\psi = e$ in (9) and using (10), we obtain

$$(a\varphi \cdot b\psi)\chi\varphi = a\varphi \cdot b\psi\chi\varphi$$

or, after replacing $a\varphi$ by a and $b\psi$ by b,

$$(a \cdot b)\chi\varphi = a \cdot b\chi\varphi \tag{12}$$

for all $a, b \in G$.

Using (11), (12), and (10) we arrive at the following equation: for arbitrary $a, b \in G$,

$$(a \circ b)\psi\chi\varphi = (a\varphi \cdot b\psi)\chi\psi\chi\varphi = (a\varphi\chi\psi \cdot b\psi)\chi\varphi = a\psi\chi\varphi \cdot b\psi\chi\varphi.$$

This equation shows that the product $\psi\chi\varphi$, which is a one-to-one mapping of G onto itself, is an isomorphism between the given semigroup and groupoid. //

A consequence of this theorem is the theorem of Albert (*Trans. Amer. Math. Soc.*, vol. 54 (1943), pp. 507-19): *If a loop is isotopic to a group, then the two are isomorphic.* From this it follows, in particular, that isotopic groups are also isomorphic, so that there is no point in employing isotopy in the theory of groups.

7. The concept of isotopy can be extended to the theory of non-associative rings. For when we discuss rings with one and the same additive group G, we can define *isotopy of rings* by the equation (3) as before, but we shall assume that the mappings φ, ψ, and χ are isomorphic mappings of the group G onto itself (in other words, automorphisms of G in the sense of § 19.1). From the results obtained above and the method used in proving them we can now derive the following results:

Every ring that is isotopic to a quasifield is itself a quasifield.

For suppose that on the additive abelian group G there are given a ring with multiplication $a \circ b$ and a quasifield with multiplication $a \cdot b$ isotopic to it. If $a \circ b = 0$, then

$$(a\varphi \cdot b\psi)\chi = 0,$$

and since the zero element remains fixed under the isomorphism χ of G, we have

$$a\varphi \cdot b\psi = 0.$$

Therefore one of the factors is zero: for example, $a\varphi = 0$; hence $a = 0$, so that the given ring does not contain divisors of zero. Thus, on the set $G \setminus 0$ there are given a groupoid and a quasigroup; and since φ, ψ, and χ can be regarded as one-to-one mappings of this set onto itself, it remains to employ § 12.4. //

Every quasifield is isotopic to a (non-associative) field.

For if a quasifield K with multiplication $a \cdot b$ is given, then by § 12.5 the multiplicative quasigroup of its non-zero elements is isotopic to a loop with multiplication $a \circ b$ by virtue of (8), where the one-to-one mappings φ and ψ of $K \setminus 0$ are defined by the equations (7). In addition, we set $0\varphi = 0\psi = 0$, which is consistent with (7). *φ and ψ are now isomorphic mappings of the additive group of K onto itself.* Hence, from

$$(a\varphi + b\varphi) \cdot f = a\varphi \cdot f + b\varphi \cdot f = a + b$$

it follows that

$$(a + b)\varphi = a\varphi + b\varphi.$$

By setting, in addition, $a \circ 0 = 0 \circ a = 0$ for all $a \in K$, we define the $\circ$-multiplication for all elements of K, the unit element e of the loop constructed above being also the unit element for this multiplication. It remains to prove that *this multiplication is distributive with respect to addition.* Thus, by (8),

$$(a + b) \circ c = (a + b)\varphi \cdot c\psi = (a\varphi + b\varphi) \cdot c\psi = \\ = a\varphi \cdot c\psi + b\varphi \cdot c\psi = a \circ c + b \circ c. \text{ //}$$

If a ring with unit element is isotopic to an associative ring, then the two rings are isomorphic.

For suppose that on an additive abelian group G there are given a ring with multiplication $a \cdot b$ having a unit element e and an associative ring with multiplication $a \circ b$, where the equation (3) holds for arbitrary $a, b \in G$, φ, ψ, and χ being isomorphic mappings of G onto itself. Then the product $\psi\chi\varphi$, which is obviously an isomorphism of G onto itself, is an isomorphism between the given rings, by what was proved in § 12.6. //

❖ An associative ring with unit element may have a non-associative isotope. There exist associative rings without unit elements that are isotopic, but not isomorphic [A. A. Albert, *Ann. of Math.*, vol. 43 (1942), pp. 685-707]. ❖

§ 13. Normal Subgroups, Ideals

1. Let G be a group and H one of its subgroups. If a is an arbitrary element of G, then the collection aH of all possible products of the form ah, where h ranges over the subgroup H, is called the *left coset* of H in G determined by a. Clearly, $a \in aH$, since H contains the unit element.

If an element b is contained in the coset aH, then $bH = aH$, that is, *every left coset of H in G is determined by each of its elements.* For if $b = ah_0$ ($h_0 \in H$) then for arbitrary $h', h'' \in H$

$$bh' = a(h_0 h') \qquad \text{and} \qquad ah'' = b(h_0^{-1} h''),$$

so that $bH \subseteq aH$ and $aH \subseteq bH$.

Hence it follows that *two arbitrary left cosets of H in G either coincide or have an empty intersection.* Thus we obtain a partition of G into disjoint left cosets of H. It is called the *left decomposition* of G with respect to the subgroup H. One of the classes of this decomposition is H itself: if $a \in H$, then $aH = H$.

2. Similarly, we can obtain the *right decomposition* of G with respect to H, consisting of the *right cosets* Ha $(a \in G)$. In the non-commutative case, the right decomposition of G with respect to H can actually differ from the left decomposition. However, both consist of the same number of classes: the mapping that carries each element a of G into a^{-1} is a one-to-one mapping of G onto itself carrying the left coset aH into the right coset Ha^{-1}, and vice versa.

The number of cosets in each of the two decompositions of G with respect to H is called the *index* of H in G, provided it is finite.

3. If G is a finite group of order n and H a subgroup of G of order k and index j, then every coset aH consists of exactly k elements, so that

$$n = kj.$$

This yields the following theorem.

LAGRANGE'S THEOREM: *The order and the index of every subgroup of a finite group are divisors of the order of the group.*

By § 9.4, it follows from this theorem that *the order of every element of a finite group is a divisor of the order of the group.* It also follows from Lagrange's Theorem that *every finite group whose order is a prime number is cyclic.* For the group must coincide with the cyclic subgroup generated by any of its elements, other than 1.

4. A subgroup H is called *normal* (or *invariant*) in G if the left decomposition of G with respect to H coincides with the right decomposition, i.e., if for every $a \in G$ we have the following equation (which is to be understood as stating the equality of the two subsets of G):

$$aH = Ha. \tag{1}$$

We can therefore speak simply of *the decomposition of G with respect to the normal subgroup H.*

Let us indicate some other definitions of a normal subgroup that are equivalent to the one given above.

Two elements x and y of a group G are called *conjugate* in G if there exists an element a in G such that

$$y = a^{-1}xa,$$

or, as we say, y is obtained from x by means of *transformation* by a.

A subgroup H of a group G is normal in G if and only if all the elements that are conjugate in G to each of the elements of H also belong to H.

For if H is normal in G and if h is an arbitrary element of H and a an arbitrary element of G, then by (1), H contains an element h' such that

$$ah' = ha, \tag{2}$$

and hence

$$a^{-1}ha = h' \in H. \tag{3}$$

Conversely, if for arbitrary $h \in H$ and $a \in G$ there exists an element h' in H satisfying the equation (3) and therefore the equation (2), then $Ha \subseteq aH$. This inclusion holds for all a and, in particular, for a^{-1}, giving $Ha^{-1} \subseteq a^{-1}H$, from which we have $aH \subseteq Ha$, so that in fact the equation (1) holds. //

Two subgroups U and V of a group G are called *conjugate* in G if G contains an element a that transforms U into V, i.e.,

$$a^{-1}Ua = V.$$

Note that for every subgroup U and every $a \in G$ the set $a^{-1}Ua$ is a subgroup isomorphic to U. For if $u_1, u_2 \in U$, then

$$(a^{-1}u_1a)(a^{-1}u_2a) = a^{-1}(u_1u_2)a;$$

on the other hand, it follows from

$$a^{-1}u_1a = a^{-1}u_2a$$

that $u_1 = u_2$, so that the mapping $u \to a^{-1}ua$ $(u \in U)$ is an isomorphism of U onto $a^{-1}Ua$ and consequently $a^{-1}Ua$ is a subgroup.

Every normal subgroup of a group G coincides with all its conjugate subgroups in G. Conversely, if a subgroup H of a group G contains all the subgroups conjugate to it in G, then it is normal.

For (1) implies that

$$a^{-1}Ha = H, \tag{4}$$

which proves the first statement of the theorem. On the other hand, if

$$a^{-1}Ha \subset H,$$

where the inclusion is strict, then

$$H \subset aHa^{-1} = (a^{-1})^{-1} Ha^{-1}.$$

Thus, if H must contain *all* subgroups conjugate to it, then it necessarily coincides with them, so that for every $a \in G$ we have the equation (4) and therefore the equation (1). //

5. *The intersection of an arbitrary set of normal subgroups of a group G is itself a normal subgroup of G.*

For if D is the intersection of the normal subgroups A_i $(i \in I)$, then D is a subgroup of G (see § 9.6) and furthermore, every element that is conjugate to an element of D is contained in each of the normal subgroups A_i and therefore belongs to D. //

The subgroup generated by an arbitrary system of normal subgroups of a group G (see § 9.8) *is itself normal in G.*

For let A_i $(i \in I)$ be normal subgroups and let B be the subgroup generated by them. As was shown in § 9.7, every element b of B can be written in the form

$$b = a_1 a_2 \ldots a_n,$$

where $a_j \in A_{i_j}$ $(j = 1, 2, \ldots, n)$. If g is an arbitrary element of G, then

$$g^{-1} b g = (g^{-1} a_1 g)(g^{-1} a_2 g) \ldots (g^{-1} a_n g),$$

and since

$$g^{-1} a_j g \in A_{i_j} \qquad (j = 1, 2, \ldots, n),$$

we have

$$g^{-1} b g \in B.$$

6. In an abelian group all subgroups, of course, are normal. In every group G the trivial subgroup E and the group G itself are normal subgroups. If a group has no normal subgroups but these two, it is called *simple*.

The finite cyclic groups of prime orders constitute the only simple abelian groups.

For if an abelian group G is simple, then it cannot have any non-trivial subgroups at all, and since it must contain cyclic subgroups, it is itself cyclic. But an infinite cyclic group has non-trivial subgroups: the additive group of integers contains the group of even integers as a proper subgroup. Furthermore, if a group $G = \{a\}$ is cyclic of finite order n, then for a composite number n $(n = kl)$ the subgroup $\{a^k\}$ is of order l, so that it is distinct from G and from E. But if $G = \{a\}$ is of prime order p, then it cannot have non-trivial subgroups, as we see from Lagrange's Theorem. //

7. Non-commutative simple groups exist, finite as well as infinite. For example, let us consider the subset A_n of even permutations in the symmetric group S_n of degree n (see § 13.8). In textbooks on algebra it is proved that the parity of a permutation of degree n is the same as the parity of the number of factors in any decomposition of the permutation into the product of transpositions. Hence it follows that the product of even permutations of degree n is always even and that the inverse of an even permutation is itself even, and we see from this that A_n *is a subgroup of* S_n.

The group A_n is called the *alternating group of degree n* and is of order $n!/2$. Note that it is not merely a subgroup, but even a normal subgroup of S_n: the index of A_n in S_n is 2, so that both of the decompositions of S_n with respect to A_n coincide, inasmuch as they consist of two classes, the subgroup A_n itself and the class of odd permutations.

The group A_3 of order 3 is a cyclic group of prime order, so that it is simple. The group A_4 is not simple: it is easy to verify that the even permutations[1]

$$(12)(34),\ (13)(24),\ (14)(23)$$

together with the unit element form a normal subgroup of this group.

The alternating group A_n *of degree n for* $n \geqq 5$ *is simple.*

Let us show first of all that the cycles (ijk) of length 3, which are even permutations, form a system of generators for A_n (see § 9.9). For, every even permutation splits into the product of an even number of transpositions and therefore into a product of cycles of length 3, because

$$(ij)(ik) = (ijk),$$
$$(ij)(kl) = (ijk)(ilk).$$

[1] We recall that in the expression of a permutation by means of cycles we write next to each symbol the one into which it is carried by the permutation in question. A cycle closes when we reach the symbol that goes over into the first symbol of the cycle. The symbols that remain fixed under our permutation can be omitted in the decomposition into cycles. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 6 & 1 & 3 & 5 & 2 \end{pmatrix} = (143)\ (26)$$

On the other hand,

$$(12)\ (34) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$

provided of course that it is known to be a permutation of degree 4. We also recall that the parity of a permutation is the same as the parity of the difference between the numbers of symbols actually permuted and the numbers of cycles in its decomposition.

Next, let us show that every normal subgroup H of A_n containing at least one cycle (ijk) of length 3 also contains every other cycle $(i'j'k')$ of length 3 and hence, as was shown above, coincides with A_n. For if the symbols l and m are different from i, j, and k, then the permutation of degree n

$$\alpha = \begin{pmatrix} i & j & k & l & m & \dots \\ i' & j' & k' & l' & m' & \dots \end{pmatrix},$$

which can be made even by transposing, if necessary, the symbols l' and m' in the bottom row, is such that

$$\alpha^{-1}(ijk)\alpha = (i'j'k').$$

It remains to show that every normal subgroup H of A_n ($n \geqq 5$), other than E, contains at least one cycle of length 3. In H we choose an element α other than 1 that leaves the largest possible number of symbols fixed. If α is not a cycle of length 3 then it is of one of the following two forms:

1) The decomposition of α into cycles contains at least one cycle whose length is not less than 3:

$$\alpha = (ijk \dots) \cdots.$$

Here α, being an even permutation, cannot be a cycle $(ijkl)$ of length 4, so that α permutes at least two symbols different from i, j, and k; let these be l and m.

2) α splits into the product of independent cycles of length 2 of which there are at least two:

$$\alpha = (ij)(kl) \cdots.$$

In this case, since $n \geqq 5$, we denote by m an arbitrary symbol different from i, j, k, and l.

Let $\beta = (klm)$. Then $\beta^{-1}\alpha\beta \in H$, since $\beta \in A_n$, and therefore

$$\gamma = \beta^{-1}\alpha\beta\alpha^{-1} \in H.$$

The permutation γ is different from 1, since in case 1) the permutation $\beta^{-1}\alpha\beta$ carries the symbol j not into k, as α does, but into l; and in case 2) $\beta^{-1}\alpha\beta$ carries the symbol l into m and not into k, as α does.

On the other hand, the permutation γ leaves fixed, in the first case, all the symbols that are left in place by α, because they are all distinct from k, l, and m, and in the second case, all the symbols left in place by α, except possibly m. However, it is easy to verify that, in the first case, γ leaves the additional symbol i in place and, in the second case, the symbols i and j. Thus, in both cases γ leaves more symbols fixed than α does. This contradiction to the choice of α shows that α must be a cycle of length 3. //

The alternating groups A_n $(n \geqq 5)$ by no means exhaust all the finite non-commutative simple groups.

8. In the next chapter, it will be shown that the role played by normal subgroups in the theory of groups is taken, in the theory of rings, by the *ideals*.

A subset A of an arbitrary ring R is called an *ideal* in R if it is a subgroup of the additive group of R—for this purpose it is sufficient that the difference of any two elements of A belong to A—and if, furthermore, for arbitrary elements $a \in A$ and $r \in R$ both the products ar and ra are contained in A.

From this definition it follows that every ideal of a ring R is a subring of R. It is also easy to see that the intersection of an arbitrary system of ideals of a ring R is itself an ideal. Further, *the subgroup of the additive group of a ring R generated by a given system of ideals A_i $(i \in I)$ is itself an ideal*: this subgroup consists of all possible finite sums of elements from the ideals A_i, and the multiplication of such a sum on the left or on the right by an arbitrary element of R leads to a sum of the same form.

In § 16.2 we gave a description of all subgroups of the additive group of integers. *These subgroups (including the zero group) constitute the only ideals in the ring of integers,* because the product of a number that is a multiple of n by an arbitrary integer is itself a multiple of n.

9. Among the ideals of an arbitrary ring R is R itself and the *zero ideal* O, consisting of the zero element only. A ring that contains no ideals except these two is called *simple*.

Every field and, more generally, every division ring (see § 12.1) *is a simple ring.*

For let K be a division ring. If A is an ideal in K, other than O, and if $a \in A$, $a \neq 0$, then every element b of R must belong to A, because the equations (1) of § 12.1 are solvable in K, and so $A = R$.

The full matrix ring K_n of every order n over an arbitrary division ring K is a simple ring.

For let A be a non-zero ideal in K_n. If $a = (a_{ij})$ is a non-zero matrix belonging to A, then, say, $a_{kl} \neq 0$. Let $\bar{x}_{ij}$ denote the matrix of K_n in which the entry at the position (i, j) is the element x of K, whereas all the remaining entries are zeros. Further, let b be an arbitrary element of K and s and t arbitrary indices $(1 \leqq s, t \leqq n)$. Since the equations (1) of § 12.1 are solvable in K, there exist elements x and y in K for which

$$y(a_{kl}x) = b.$$

By applying the multiplication rule for matrices, we see that

$$\bar{y}_{sk}(a\bar{x}_{lt}) = \bar{b}_{st}.$$

Thus, all the matrices of the form $\bar{b}_{st}$ belong to A, and since every matrix of K_n can be represented as a sum of matrices of this form, we see that $A = K_n$, as required. //

10. If the condition is dropped from the definition of an ideal that for every a of A and for every r of R both of the products ar and ra must belong to A and if it is required instead only that this holds for the product ar alone or, alternatively, for the product ra alone, then we arrive at the concept of a one-sided ideal—a *right ideal* if $ar \in A$ for all $a \in A$ and $r \in R$, and a *left ideal* if $ra \in A$. For commutative rings as well as for anticommutative rings, and in particular for Lie and Jordan rings (see §§ 8.3 and 8.4), every one-sided ideal is, of course, an ideal (or, as it is sometimes called, a *two-sided ideal*).

In our proof above that every division ring is simple we only had to make use of the fact that one of the two equations (1) of § 12.1 is solvable; hence we may claim that *no division ring K can contain one-sided ideals other than K and O.*

If an associative ring R contains no one-sided ideals except R and O and is not a ring with trivial multiplication (see § 8.2), *then it is a field.*

For let us define a *left annihilator* of a ring R as an element a for which $ar = 0$ for all $r \in R$. The left annihilators in an associative ring R form an ideal, and even a two-sided ideal. In our case, this is the zero ideal: if it were R, then R would be a zero ring. Thus, for every non-zero element a the set aR of all possible products of the form ar ($r \in R$), being a right ideal in an associative ring, is different from 0, and so $aR = R$. This implies that for every element $b \in R$ the equation $ax = b$ is solvable in R. Similarly, it can be shown that every equation of the form $ya = b$, $a \neq 0$, is solvable in R, so that by § 8.10 the ring R is a field. //

From what we have said it follows easily that *a non-zero associative-commutative ring is simple if and only if it is a field.* On the other hand, since in a zero ring every subgroup of the additive group is an ideal, we see by § 13.6 that *simple zero rings are zero rings on cyclic additive groups of prime order.*

11. The definitions of an ideal and a one-sided ideal carry over readily to the case of an arbitrary groupoid: we need only eliminate the references to the additive group from these definitions. Every ideal, left ideal, and right ideal of a ring is obviously an ideal, left ideal, and right ideal, respectively, of the multiplicative groupoid of the ring. The converse, however, is not true: in the ring of integers the set of numbers that are multiples of

at least one of the numbers 2 and 3 is not an ideal, although it is an ideal in the multiplicative semigroup of the ring.

§ 14. Gaussian Semigroups

1. It is well known that in the ring of integers C and in the ring of polynomials $P[x]$ over an arbitrary commutative field P the theories of divisibility are completely parallel. In this and the following sections we shall clarify the nature of this parallelism.

Let G be an abelian semigroup with unit element 1 satisfying the cancellation law (see § 11.1). A *divisor of unity* is an element ε for which G contains an inverse element ε^{-1},

$$\varepsilon\varepsilon^{-1} = 1.$$

The divisors of unity form a subgroup of the semigroup, because the product of divisors of unity and the element inverse to a divisor of unity are themselves divisors of unity. On the other hand, *if the product of the elements* $a_1, a_2, \ldots, a_n \in G$ *is a divisor of unity,*

$$a_1 a_2 \ldots a_n = \varepsilon,$$

then each of the elements a_i $(i = 1, 2, \ldots, n)$ *is itself a divisor of unity.* In fact,

$$a_i^{-1} = a_1 \ldots a_{i-1} a_{i+1} \ldots a_n \varepsilon^{-1} \qquad (i = 1, 2, \ldots, n).$$

2. If for $a, b \in G$ there is a $c \in G$ such that $a = bc$, then b is called a *divisor* of a; we also say that a is *divisible* by b. Two elements a and b are called *associates* if each of them is a divisor of the other,

$$a = bc, \quad b = ad. \tag{1}$$

Since (1) implies that

$$a = a(cd),$$

so that, in virtue of the cancellation law,

$$cd = 1,$$

then both c and d are divisors of unity. On the other hand, for every divisor of unity ε *the elements a and* εa *are associates,* because

$$a = \varepsilon^{-1}(\varepsilon a).$$

From this it follows that *each divisor of unity is a divisor of every element of G.*

The relation of being associates is obviously an equivalence relation, so that G splits into *classes of associate elements;* one of these classes is the group of divisors of unity. On the other hand, the relation 'b divides a' is reflexive and transitive. Moreover, for arbitrary divisors of unity ε_1 and ε_2 the relation

$$a = bc$$

implies that

$$\varepsilon_1 a = (\varepsilon_2 b)(\varepsilon_1 \varepsilon_2^{-1} c),$$

i.e., $\varepsilon_2 b$ divides $\varepsilon_1 a$. Thus, we obtain a *partial order* in the set of classes of associate elements of G if for two such classes A, B we put $B \leqq A$ whenever at least one (and therefore every) element of B divides at least one (and therefore every) element of A. The class of divisors of unity is the unique minimal element of this set (see § 5.1).

3. An element $p \in G$ that is not a divisor of unity is called *irreducible* if its only divisors, apart from divisors of unity, are its associates, that is, if it always follows from

$$p = ab$$

that of the elements a, b one is a divisor of unity and the other, accordingly, an associate of p. *If p is irreducible, then, obviously, so are all its associates.* The classes of associate irreducible elements, provided such elements exist at all in G, are precisely the minimal elements of the partially ordered set of classes of associate elements, other than the class of divisors of unity.

An element $p \in G$ that is not a divisor of unity is called a *prime* if a product ab is divisible by p only when at least one of the elements a, b is divisible by p. *Every element that is an associate of a prime element p is obviously prime.* By mathematical induction on n we can also prove that *if a product $a_1 a_2 \ldots a_n$ is divisible by a prime element p, then at least one of the elements a_i $(i = 1, 2, \ldots, n)$ is divisible by p.*

Every prime element p is irreducible.

For if

$$p = ab,$$

then, since p is prime, at least one of the factors, say a, is divisible by p. However, p in turn is divisible by a; that is, p and a are associates, and therefore b is a divisor of unity. //

Note that the converse statement does not hold in general.

4. If $a, b \in G$, then an element $d \in G$ is called *a greatest common divisor* of a and b and is denoted by

$$d = (a, b) \tag{2}$$

if d is a common divisor of a and b and if d is itself divisible by every other common divisor of a and b.

If we also have $d' = (a, b)$, then d and d' must divide each other, i.e., they are associates. On the other hand, if $d = (a, b)$ and if $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are arbitrary divisors of unity, then we also have

$$\varepsilon_1 d = (\varepsilon_2 a, \varepsilon_3 b),$$

that is, when all the elements occurring in (2) are replaced by arbitrary elements associated with them, the equation remains valid. In other words, if A, B, and D are the classes of elements associated with a, b, and d, respectively, then (2) can be rewritten in the form

$$D = (A, B).$$

The class D is the unique maximal element (see § 5.5) in the partially ordered set of all those classes X for which $X \leqq A$ and $X \leqq B$, and the symbol (A, B) can be regarded as the notation for this class.

Of course, two elements $a, b \in G$ need not, in general, have a greatest common divisor. However, *if a is divisible by b, then (a, b) exists and*

$$(a, b) = b.$$

5. *If every pair of elements $a, b \in G$ has a greatest common divisor, then every irreducible element of G is prime.*

Let us show first of all that under our assumptions we have for arbitrary $a, b, c \in G$

$$(ac, bc) = (a, b)c. \tag{3}$$

For since a is divisible by (a, b), ac is divisible by $(a, b)c$. Similarly, bc is divisible by $(a, b)c$, so that the left-hand side of (3) is divisible by the right-hand side:

$$(ac, bc) = (a, b)cd. \tag{4}$$

Hence

$$ac = (ac, bc)u = (a, b)cdu,$$

so that by the cancellation law,

$$a = (a, b)du. \tag{5}$$

Similarly,

$$b = (a, b)dv. \tag{6}$$

From (5) and (6) it follows that $(a, b)d$ is a common divisor of a and b, and as it is divisible by (a, b), the element d must be a divisor of unity. We have thus shown that we can go from (4) to (3).

Next let us show that for any three elements $a, b, c \in G$ we have

$$((a, b), c) = (a, (b, c)). \tag{7}$$

For the element $((a, b), c)$ is a common divisor of (a, b) and c and therefore a common divisor of a, b, and c. On the other hand, every common divisor of these three elements is a common divisor of (a, b) and c, and is therefore a divisor of $((a, b), c)$. The same statement can be made of the element $(a, (b, c))$, so that the left-hand and right-hand sides of (7) are associates, and this is essentially the content of (7).

Now let us call two elements $a, b \in G$ *relatively prime* if their greatest common divisor is a divisor of unity; this can be expressed by

$$(a, b) = 1.$$

If $(a, b) = 1$ *and* $(a, c) = 1$, *then we also have* $(a, bc) = 1$.

For since a is a divisor of ac, we have

$$(a, ac) = a.$$

Further, from $(a, b) = 1$ it follows by (3) that

$$(ac, bc) = c.$$

Therefore, by (7),

$$(a, bc) = ((a, ac), bc) = (a, (ac, bc)) = (a, c) = 1.$$

Let us now proceed to the proof of the theorem itself. Suppose that the product ab is divisible by p, where p is an irreducible element of G. If neither a nor b is divisible by p, then in virtue of the irreducibility of p the only common divisors of a and p (and also, of b and p) are the divisors of unity, that is, a and p (and also, b and p) are relatively prime:

$$(a, p) = 1, \quad (b, p) = 1.$$

But then, by what we have shown above, we also have

$$(ab, p) = 1,$$

although by assumption ab is divisible by p, and p is not a divisor of unity. This contradiction proves that p is a prime. //

6. Let G, as before, be an abelian semigroup with unit element 1 satisfying the cancellation law. If

$$a = b_1 b_2 \dots b_k \tag{8}$$

and if the elements $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k$ are such that

$$\varepsilon_1\varepsilon_2 \dots \varepsilon_k = 1$$

—all these elements are therefore divisors of unity—then we also have

$$a = (\varepsilon_1 b_1)(\varepsilon_2 b_2) \dots (\varepsilon_k b_k). \tag{9}$$

Two decompositions of an element a into the product of several elements, (8) and

$$a = c_1 c_2 \dots c_l, \tag{10}$$

are called *associate decompositions* if $l = k$ and if, after a suitable change in the numbering of the factors in the second decomposition, the elements b_i and c_i $(i = 1, 2, \dots, k)$ are associates, i.e., if (10) is of the form (9).

An abelian semigroup G with unit element satisfying the cancellation law is called a *Gaussian semigroup* if every element a of G that is not a divisor of unity can be decomposed into the product of irreducible elements and if any two such decompositions of a are associates.

The Gaussian semigroups comprise, in particular, all the abelian groups: these do not contain elements that are not divisors of unity. The multiplicative semigroup of non-zero integers is also Gaussian. The divisors of unity in this semigroup are the numbers 1 and -1, and the irreducible elements are all the prime numbers, taken with the sign $+$ or $-$. The multiplicative semigroup of the non-zero polynomials in an indeterminate x with coefficients from a commutative field P is also a Gaussian semigroup; its divisors of unity are all the non-zero elements of P, and the irreducible elements are exactly the irreducible polynomials.

7. *An abelian semigroup G with unit element satisfying the cancellation law is Gaussian if and only if it satisfies* (α) *and, in addition,* (β') *and* (β'') *of the following three conditions*:

(α) *The partially ordered set of classes of associate elements of G* (see § 14.2) *satisfies the minimal condition* (see § 5.1);

(β') *Any two elements of G have a greatest common divisor;*

(β'') *Every irreducible element of G is prime.*

Since, by § 14.5, (β') implies (β''), this theorem reduces to the two statements to be proved below.

8. *Every Gaussian semigroup G satisfies the conditions* (α) *and* (β').

For let us define the *length* of an element a of a Gaussian semigroup G as the number of factors in any decomposition of a into the product of irreducible factors. If $a = bc$ and b is a *proper* divisor of a, i.e., if neither b nor c are divisors of unity, then we obtain a decomposition of a into the product of irreducible factors by multiplying the decompositions of b and

c of the same kind. Thus, the length of a is equal to the sum of the lengths of b and c, so that the length of a proper divisor of a is strictly less than that of a. Hence it follows that every sequence

$$a_1, a_2, \ldots, a_n, \ldots$$

of elements of G for which a_{n+1} is a proper divisor of a_n $(n = 1, 2, \ldots)$ must break off after a finite number of terms. This proves the validity of (α).

On the other hand, suppose that $a, b \in G$ and that

$$p_1, p_2, \ldots, p_n \tag{11}$$

is a set of irreducible elements such that every irreducible divisor of a as well as of b is an associate of one and only one of the elements of (11); it follows that among the elements (11) there are no associates. The elements a and b can now be written in the form

$$a = \varepsilon p_1^{k_1} p_2^{k_2} \ldots p_n^{k_n}, \qquad b = \varepsilon' p_1^{l_1} p_2^{l_2} \ldots p_n^{l_n},$$

where ε and ε' are divisors of unity and where some of the exponents $k_1, k_2, \ldots, k_n, l_1, l_2, \ldots, l_n$ may be equal to zero. Since every divisor c of a can be written in the form

$$c = \varepsilon'' p_1^{s_1} p_2^{s_2} \ldots p_n^{s_n},$$

where ε'' is a divisor of unity and $0 \leqq s_i \leqq k_i$ $(i = 1, 2, \ldots, n)$ and since a similar statement is true of the divisors of b, a greatest common divisor of a and b is the element

$$\alpha = p_1^{m_1} p_2^{m_2} \ldots p_n^{m_n},$$

where $m_i = \min(k_i, l_i)$ $(i = 1, 2, \ldots, n)$. Thus, the condition (β') is satisfied. //

9. *Every abelian semigroup G with unit element satisfying the cancellation law and the conditions (α) and (β'') is Gaussian.*

We have to show that every element $a \in G$ that is not a divisor of unity has a decomposition into the product of irreducible elements and that this decomposition is unique to within associates. First we observe that if this statement holds for a, then it also holds for every associate of a. This enables us, by (α) and § 5.1, to prove our statement by induction on the partially ordered set of classes of associate elements of G.

Our statement necessarily holds for irreducible elements, that is, for the elements of the minimal classes other than the class of divisors of unity.

We may therefore assume that everything is proved for all the proper divisors of a. If a is not an irreducible element, then

$$a = bc,$$

where b and c are proper divisors of a. By the induction hypothesis,

$$b = p_1 p_2 \dots p_k, \qquad c = p_{k+1} p_{k+2} \dots p_n,$$

where all the p_i $(i = 1, 2, \dots, n)$ are irreducible. Hence

$$a = p_1 p_2 \dots p_k p_{k+1} p_{k+2} \dots p_n. \tag{12}$$

Suppose next that

$$a = q_1 q_2 \dots q_s \tag{13}$$

is any other decomposition of a into the product of irreducible factors. Since by (β'') the irreducible element q_1 is a prime, we know by § 14.3 that at least one of the elements p_i $(i = 1, 2, \dots, n)$ must be divisible by q_1. Let this be p_1. It is irreducible, however, and therefore the elements p_1 and q_1 are associates,

$$p_1 = \varepsilon q_1. \tag{14}$$

Hence by the cancellation law,

$$(\varepsilon p_2) p_3 \dots p_n = q_2 q_3 \dots q_s. \tag{15}$$

The left-hand and right-hand sides of (15) are decompositions of a proper divisor of a into irreducible factors and so, by the induction hypothesis, they are associates. But then it follows from this and from (14) that the decompositions (12) and (13) of a are associate decompositions; in fact, since the elements εp_2 and q_2, say, are associates, so are p_2 and q_2. //

§ 15. Gaussian Rings

1. We recall that if R is an integral domain (see § 8.7), then the set $R \setminus 0$, under the operation of multiplication, is an abelian semigroup satisfying the cancellation law.

An integral domain R with unit element is called a *Gaussian ring* if the multiplicative semigroup of the non-zero elements of R is a Gaussian semigroup.

The Gaussian rings include, trivially, all the commutative fields. However, not every integral domain with unit element is a Gaussian ring.

❖ The set of complex numbers of the form $a + b\sqrt{-3}$, where a and b are arbitrary integers, is an integral domain with unit element, but not a Gaussian ring. For example, in this ring

$$4 = 2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3})$$

are two non-associate decompositions of the number 4 into the product of irreducible factors. ❖

2. Let R be an integral domain with unit element. If $a \in R$, then the set

$$(a) = aR,$$

that is, the collection of elements of the form ar $(r \in R)$ is an ideal in R (see § 13.8). This is the *principal ideal* generated by a. Clearly a is contained in (a), because $a = a \cdot 1$. If all the ideals of R are principal ideals, that is, if every ideal of R can be generated by a single element, then R is called a *principal ideal ring*.

Every principal ideal ring R is a Gaussian ring.

By § 14.7, it is sufficient to prove that the multiplicative semigroup of the non-zero elements of R satisfies the conditions (α) and (β').

If a is divisible by b, that is, $a = bc$, then $a \in (b)$ and therefore $(a) \subseteq (b)$; the converse is also true. Thus, if in a sequence of non-zero elements

$$a_1, a_2, \ldots, a_n, \ldots \tag{1}$$

each element a_n is divisible by a_{n+1} $(n = 1, 2, \ldots)$, then the corresponding principal ideals form an ascending sequence

$$(a_1) \subseteq (a_2) \subseteq \ldots \subseteq (a_n) \subseteq \ldots. \tag{2}$$

It is easy to verify that the set-theoretical union of this ascending sequence of ideals is itself an ideal in R, so that by virtue of our assumptions on R it is a principal ideal; we denote it by (b). The element b, which belongs to the union of the ascending sequence (2), must be contained in some ideal (a_n) and therefore in all the ideals (a_i) for $i \geqq n$. For all these i, consequently, we have $(b) \subseteq (a_i)$, and this, together with $(a_i) \subseteq (b)$, yields $(a_i) = (b)$. Thus, the elements a_i and b divide each other, i.e., they are associates. This shows that in (1) all the elements a_i for $i \geqq n$ are associates of each other, that is, the condition (α) holds.

On the other hand, if $a, b \in R$, then the collection of elements of the form $ar + bs$, where r and s range independently over the whole ring R, is an ideal and therefore a principal ideal; we denote it by (d). Then

$$(d) \supseteq (a), \qquad (d) \supseteq (b),$$

since

$$a = a \cdot 1 + b \cdot 0 \in (d), \quad b = a \cdot 0 + b \cdot 1 \in (d),$$

and so d is a common divisor of a and b. If c is an arbitrary common divisor of a and b, then

$$(c) \supseteq (a), \qquad (c) \supseteq (b),$$

so that (c) contains every element of the form $ar + bs$ or, in other words, $(c) \supseteq (d)$, from which it follows that c is also a divisor of d. Thus we see that d is a greatest common divisor of a and b. This shows that the condition (β') also holds. //

3. An integral domain R with unit element is called a *euclidean ring* if to every element $a \in R$ other than zero there can be assigned a non-negative integer $n(a)$ such that the following conditions are satisfied: for arbitrary elements $a, b \in R$, where $b \neq 0$, we can find elements q and r in R for which

$$a = bq + r,$$

and either $r = 0$ or $n(r) < n(b)$.

Every euclidean ring is a principal ideal ring and therefore a Gaussian ring.

For let A be an ideal in a euclidean ring R. If $A = O$, then $A = (0)$. But if $A \neq O$, then let a_0 be one of the non-zero elements of A for which $n(a_0) \leqq n(a)$, whatever the non-zero element $a \in A$. By assumption we can then find, for an arbitrary $a \in A$, elements q and r in R such that

$$a = a_0 q + r,$$

where $n(r) < n(a_0)$ if $r \neq 0$. However, in that case

$$r = a - a_0 q \in A,$$

and we arrive at a contradiction to the choice of a_0. Therefore $r = 0$, that is, $a = a_0 q$, and this shows that A is the principal ideal generated by a_0. //

The euclidean rings include the ring of integers C—the role of $n(a)$ is taken here by the absolute value $|a|$ of the number a—*as well as the ring of polynomials $P[x]$ over a commutative field P*—here the degree of the polynomial takes the role of $n(a)$. *Therefore C and $P[x]$ are principal ideal rings and Gaussian rings.*

❖ The euclidean algorithm for the computation of the greatest common divisor of two elements, which is familiar to the reader, can be used in every euclidean ring.

There exist principal ideal rings that are not euclidean rings [T. S. Motzkin, *Bull. Amer. Math. Soc.*, vol. 55 (1949), pp. 1142-46]. ❖

4. We conclude the present section by proving the following theorem:

If R is a Gaussian ring, then the polynomial ring $R[x]$ is also Gaussian.

We know from § 8.7 that $R[x]$ is in any case an integral domain with a unit element. We have to show, by § 14.7, that the conditions (α) and (β'') of that section hold in the multiplicative semigroup of the non-zero elements of $R[x]$.

We begin with the remark that *the only divisors of unity of $R[x]$ are those of R.* Next, we mention that, since the non-zero elements of the Gaussian ring R satisfy the condition (β') of § 14.7, by (7) of § 14.5, *we can speak of the greatest common divisor, uniquely determined to within associates, of an arbitrary finite system of non-zero elements of R.* Clearly, this remains valid when a number of elements equal to zero are added to the system in question.

A polynomial $\varphi(x)$ is called *primitive* if the greatest common divisor of the system of its coefficients is a divisor of unity and can therefore be taken to be 1. A polynomial that is an associate of a primitive polynomial is obviously itself primitive. The only primitive polynomials of degree zero are the divisors of unity.

If $f(x)$ is an arbitrary non-zero polynomial and a is the greatest common divisor of its coefficients, then

$$f(x) = a\varphi(x), \tag{3}$$

where $\varphi(x)$ is a primitive polynomial. If we also have the equation

$$f(x) = b\psi(x),$$

where $b \in R$ and $\psi(x)$ is a primitive polynomial, then b, being a divisor of all the coefficients of $f(x)$, must also divide a:

$$a = bc.$$

As there are no divisors of zero in $R[x]$, we see that

$$\psi(x) = c\varphi(x);$$

and since $\psi(x)$ is primitive, c must be a divisor of unity.

Hence, *in the expression* (3) *for $f(x)$ both the elements a of R and the primitive polynomial $\varphi(x)$ are uniquely determined to within factors that are divisors of unity, i.e., to within associates.* We shall call (3) the *canonical expression* of $f(x)$.

5. GAUSS'S LEMMA: *The product of primitive polynomials is itself primitive.*

For suppose that

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_ix^{k-i} + \ldots + a_k,$$
$$g(x) = b_0x^l + b_1x^{l-1} + \ldots + b_jx^{l-j} + \ldots + b_l$$

are two primitive polynomials and that their product

$$f(x)g(x) = c_0x^{k+l} + c_1x^{k+l-1} + \ldots + c_{i+j}x^{(k+l)-(i+j)} + \ldots + c_{k+l}$$

is not primitive. Since R is a Gaussian ring, this means that the coefficients $c_0, c_1, \ldots, c_{k+l}$ have a common irreducible divisor p. Since $f(x)$ and $g(x)$ are primitive polynomials, p cannot divide all the coefficients of either of them. Of the coefficients not divisible by p, let a_i and b_j be those with the smallest index. But

$$c_{i+j} = a_ib_j + a_{i-1}b_{j+1} + a_{i-2}b_{j+2} + \ldots + a_{i+1}b_{j-1} + a_{i+2}b_{j-2} + \ldots \qquad (4)$$

Since by assumption all the coefficients $a_{i-1}, a_{i-2}, \ldots$ and $b_{j-1}, b_{j-2}, \ldots$ are divisible by p, the left-hand side of (4) is divisible by p and so are all the terms of the right-hand side except the first. Hence it follows that the product a_ib_j must be divisible by p, and inasmuch as in the Gaussian ring R the irreducible element p is prime (see § 14.7), at least one of the elements a_i, b_j must be divisible by p, in contradiction to the way they were chosen. This completes the proof of the lemma.

It follows from the lemma that *if* $f_i(x)$ $(i = 1, 2, \ldots, n)$ *are non-zero polynomials with the canonical expression*

$$f_i(x) = a_i\varphi_i(x) \qquad (i = 1, 2, \ldots, n)$$

and if

$$f(x) = \prod_{i=1}^{n} f_i(x), \quad a = \prod_{i=1}^{n} a_i, \quad \varphi(x) = \prod_{i=1}^{n} \varphi_i(x),$$

then

$$f(x) = a\varphi(x)$$

is the canonical expression for $f(x)$. Hence, *if the product of several polynomials is a primitive polynomial, then each of the factors is primitive.* We also mention that *if* $f(x)$ *and* $g(x)$ *are polynomials with the canonical expressions*

$$f(x) = a\varphi(x), \qquad g(x) = b\psi(x)$$

and if $f(x)$ *is divisible by* $g(x)$, *then* a *is divisible by* b *and* $\varphi(x)$ *by* $\psi(x)$.

6. Suppose that in $R[x]$ a sequence of non-zero polynomials

$$f_1(x), f_2(x), \ldots, f_n(x), \ldots \tag{5}$$

is given such that $f_n(x)$ is divisible by $f_{n+1}(x)$ $(n = 1, 2, \ldots)$. If

$$f_n(x) = a_n \varphi_n(x) \qquad (n = 1, 2, \ldots)$$

are the canonical expressions for these polynomials, then in the two sequences

$$a_1, a_2, \ldots, a_n, \ldots, \tag{6}$$

$$\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots, \tag{7}$$

every element is divisible by the following.

Since R is a Gaussian ring and, on the other hand, the degrees of the polynomials in (7) do not increase, we can find a natural number N with the following properties: 1) All the elements $a_n \in R$ for $n \geqq N$ are associates; 2) All the polynomials $\varphi_n(x)$ with $n \geqq N$ have one and the same degree, i.e., differ from each other by factors from R which, because the polynomials $\varphi_n(x)$ are primitive, must be divisors of unity. Hence it follows that all the polynomials $f_n(x)$ in (5) with $n \geqq N$ are associates and therefore that *the non-zero polynomials of* $R[x]$ *satisfy condition* (a).

7. Suppose now that an *irreducible* polynomial $p(x)$ divides the product $f(x)g(x)$. If

$$f(x) = a\varphi(x), \qquad g(x) = b\psi(x)$$

are the canonical expressions for $f(x)$ and $g(x)$, then

$$f(x)g(x) = (ab)[\varphi(x)\psi(x)]$$

is the canonical expression for $f(x)g(x)$. On the other hand, the polynomial $p(x)$ also has a canonical expression and, as it is irreducible, $p(x)$ must be either an irreducible element p of R or an irreducible primitive polynomial $\pi(x)$.

In the first case, p divides the product ab, and as R is a Gaussian ring, p must divide at least one of the elements a, b, say a. In that case, p is a divisor of the polynomial $aq(x) = f(x)$.

8. There remains the more difficult second case. Here the polynomial $\pi(x)$ is a divisor of the product of primitive polynomials $\varphi(x)\psi(x)$.

As we know from § 11.5, the integral domain R can be embedded in its field of fractions Q. Every polynomial in $R[x]$ is, of course, also a polynomial in $Q[x]$. On the other hand, if $q(x)$ is an arbitrary non-zero

polynomial of $Q[x]$, then we write its coefficients as fractions with a least common denominator and factor out this least common denominator; next, we take the greatest common divisor of the numerators and arrive at the expression

$$q(x) = \frac{a}{b}\varphi(x), \tag{8}$$

where $\varphi(x)$ is a primitive polynomial in $R[x]$.

If

$$q(x) = \frac{c}{d}\psi(x)$$

is another expression of the form (8) for $q(x)$, i.e., if the polynomial $\psi(x)$ belongs to $R[x]$ and is primitive, then

$$(ad)\varphi(x) = (bc)\psi(x), \tag{9}$$

so that the left-hand and right-hand sides of (9) are canonical expressions for one and the same polynomial in $R[x]$ and hence are associates. Thus, *the primitive polynomial $\varphi(x)$ is uniquely determined to within associates by $q(x)$.* We call (8) the *canonical expression* for the polynomial $q(x)$ in $Q[x]$.

If $q_1(x)$, $q_2(x)$, are non-zero polynomials in $Q[x]$ with the canonical expressions

$$q_i(x) = \frac{a_i}{b_i}\varphi_i(x) \qquad (i = 1, 2), \tag{10}$$

then by Gauss's Lemma,

$$q_1(x)q_2(x) = \frac{a_1a_2}{b_1b_2}[\varphi_1(x)\varphi_2(x)] \tag{11}$$

is the canonical expression for the product $q_1(x)q_2(x)$.

The irreducible primitive polynomial $\pi(x)$ remains irreducible in the ring $Q[x]$. For if

$$\pi(x) = q_1(x)q_2(x),$$

where $q_1(x)$, $q_2(x) \in Q[x]$ have the canonical expression (10), then by (11) $\pi(x)$ is an associate of the product $\varphi_1(x)\varphi_2(x)$. Therefore one of the factors $\varphi_i(x)$ $(i = 1, 2)$ must be of degree 0; the same is true for the corresponding $q_i(x)$.

We know from § 15.3. that $Q[x]$ is a Gaussian ring. Hence it follows, since $\pi(x)$ divides the product $\varphi(x)\psi(x)$, that one of these factors, $\varphi(x)$ say, is divisible by $\pi(x)$ *in the ring* $Q[x]$:

$$\varphi(x) = \pi(x)q(x), \qquad q(x) \in Q[x].$$

If $q(x)$ has the canonical expression

$$q(x) = \frac{c}{d}\chi(x),$$

then it follows from the primitivity of $\varphi(x)$ and $\pi(x)$ that $\varphi(x)$ is an associate of the product $\pi(x)\chi(x)$, so that $\varphi(x)$ is divisible by $\pi(x)$ *even in the ring* $R[x]$. Consequently, the polynomial $f(x) = a\varphi(x)$ is also divisible by $\pi(x)$.

Thus, *condition* (β'') *holds in* $R[x]$. The proof of the theorem of § 15.4 is now complete. //

9. It follows immediately from this theorem that *the ring of polynomials* $R[x_1, x_2, \ldots, x_n]$ *in an arbitrary finite number of indeterminates over an arbitrary Gaussian ring* R *is itself Gaussian.* In particular, the ring of polynomials $P[x_1, x_2, \ldots, x_n]$ over an arbitrary field P is Gaussian.

It is now easy to prove that *there exist Gaussian rings that are not principal ideal rings.* For example, in the ring of polynomials $R = P[x, y]$ over a field P, the set A of polynomials without a constant term is an ideal different from R. It is not, however, a principal ideal, because the only common divisors of the two polynomials x and y that lie in A are the non-zero elements of P.

§ 16. Dedekind Rings

1. As in the preceding section, we shall discuss here an integral domain R with unit element. We recall (see §§ 11.5 and 11.6) that an integral domain R is contained in a uniquely determined field of fractions.

If A and B are ideals of R, then their *product* AB is defined as the ideal generated by all possible products of the form ab, where $a \in A$, $b \in B$. It is easy to see that the ideal AB consists of precisely those elements of R that can be written in at least one way in the form

$$\sum_{i=1}^{n} a_i b_i, \qquad a_i \in A, \quad b_i \in B. \tag{1}$$

The associativity of the operations in R at once makes the multiplication of ideals associative. We can, therefore, speak of the *semigroup of ideals* of R; it is obviously a commutative semigroup.

Since every element of the form (1) also belongs to the intersection of the ideals A and B, we have

$$AB \subseteq A \cap B.$$

Furthermore, it obviously follows from $B \subseteq B'$ that

$$AB \subseteq AB' \tag{2}$$

2. An ideal C is called *prime* if the relations $xy \in C$ $(x, y \in R)$ always imply that at least one of the elements x, y is contained in C. The ideals R and O clearly satisfy this definition. *Henceforth, in speaking of prime ideals, we shall exclude these two ideals from consideration.*

C is a prime ideal if and only if for arbitrary ideals A, B it follows from $AB \subseteq C$ that either $A \subseteq C$ or $B \subseteq C$.

For suppose that $AB \subseteq C$, but $A \not\subseteq C$, $B \not\subseteq C$. Therefore, there exist elements $a \in A$, $b \in B$ that are not contained in C, although $ab \in C$. Thus, C is not a prime ideal. Conversely, if C is not a prime ideal, then there exist a and b not in C but with $ab \in C$. Going over to principal ideals, we see that $(a) \not\subseteq C$, $(b) \not\subseteq C$, but $(a)(b) \subseteq C$. //

By a *maximal ideal* of our ring R we mean one that is maximal among the proper ideals (i.e., different from R).

Every maximal ideal M is prime.

For suppose that $a \notin M$, $b \notin M$, but

$$ab \in M. \tag{3}$$

The ideal (M, a) generated by M and a coincides with R. However, it consists of elements that can be written in the form $m + xa$ $(m \in M, x \in R)$ and can therefore also be written in the form $M + Ra$. Similarly, $(M, b) = R$, and this ideal can be written in the form $M + Rb$. Therefore, by (3), we arrive at a contradiction:

$$R = RR = (M + Ra)(M + Rb) \subseteq M. \text{ //}$$

3. *In a principal ideal ring* (see § 15.2) *the prime ideals coincide with the ideals of the form (p), where p is a prime element* (see § 14.3).

For if $xy \in (p)$, then

$$xy = pz \qquad (z \in R),$$

so that by the definition of a prime element at least one of the elements x, y is divisible by p, i.e., is contained in (p). But if an element q is not a prime, then by § 14.7 it can be reduced to the form $q = ab$, where neither a nor b are divisors of unity. Hence $ab \in (q)$; however, if $a \in (q)$, for example, then

$$a = qc = abc,$$

and as there are no divisors of zero, $bc = 1$, so that b proves to be a divisor of unity. //

In a principal ideal ring every ideal other than the whole ring and zero is a product of a finite number of prime ideals.

For if an element a is not a divisor of unity and not equal to zero, then a has a decomposition into the product of prime factors,

$$a = p_1 p_2 \dots p_n,$$

so that we have

$$(a) = (p_1)(p_2) \dots (p_n). \text{ //}$$

An integral domain R with unit element is called a *Dedekind ring* if every ideal of R other than R itself and O can be represented as a product of a finite number of prime ideals. Below, we shall give a characterization of Dedekind rings.

4. Let R be an integral domain with unit element and P, its field of fractions. A subgroup A of the additive group P is called a *fractional ideal* of R if the following two conditions are satisfied: 1) If $a \in A$, $x \in R$, then $ax \in A$; 2) All the elements of A can be written as fractions with a common denominator:

$$A = \frac{1}{d} A_0, \quad A_0 \subseteq R, \quad d \in R, \quad d \neq 0. \tag{4}$$

Clearly, A_0 is an ideal in R, and the conditions (4) can be regarded as a definition of fractional ideals.

The fractional ideals include the ideals, or as we shall call them henceforth, the *integral ideals* of R. A *principal fractional ideal* is defined as a fractional ideal of the form $\frac{1}{b}(a)$, where (a) is a principal integral ideal. This principal fractional ideal contains, in particular, the element $\frac{a}{b}$ of P and is generated by this element, so that it could have been written in the form $\left(\frac{a}{b}\right)$.

Quite generally, since the intersection of an arbitrary set of fractional ideals is a fractional ideal, we can speak of the fractional ideal *generated* by a given finite set of elements of P: since all these elements can be written as fractions with a common denominator d, they all lie in the fractional ideal $\frac{1}{d} R$, so that we can speak of the intersection of all fractional ideals containing the given elements.

5. We can extend the definition of the product of ideals to fractional ideals: if

$$A = \frac{1}{c} A_0, \qquad B = \frac{1}{d} B_0,$$

then

$$AC = \frac{1}{cd} A_0 B_0,$$

so that *this product is itself a fractional ideal.* Since multiplication remains associative (and commutative) and the product of non-zero fractional ideals is itself different from zero, we can speak of the *semigroup of non-zero fractional ideals* of R. The unit element of this semigroup is R itself.

Our object is to prove the following theorem:

An integral domain R with unit element is a Dedekind ring if and only if its semigroup of non-zero fractional ideals is a group.

6. Let R, as before, be an integral domain with unit element. A fractional ideal A in R is called *invertible* if there exists another fractional ideal A^{-1} such that

$$AA^{-1} = R.$$

Every non-zero principal fractional ideal is invertible, because

$$\left(\frac{a}{b}\right)^{-1} = \left(\frac{b}{a}\right).$$

Every invertible fractional ideal is generated by a finite number of elements.

For if $AA^{-1} = R$, then the unit element has an expression of the form

$$1 = \sum_{i=1}^{n} a_i a_i' \qquad (a_i \in A, \quad a_i' \in A^{-1}, \quad i = 1, 2, \ldots, n).$$

Hence we have for every $a \in A$

$$a = \sum_{i=1}^{n} a_i (a_i' a),$$

and since $a_i' a \in R$ $(i = 1, 2, \ldots, n)$,

$$A = (a_1, a_2, \ldots, a_n).$$

If

$$B = A_1 A_2 \ldots A_n$$

and if the fractional ideal B is invertible, then each of the fractional ideals A_i $(i = 1, 2, \ldots, n)$ is also invertible, because

$$A_i \Big[\Big(\prod_{j \neq i} A_j \Big) \Big] B^{-1} = R.$$

7. *If the prime ideals $C_1, C_2, \ldots, C_n$, regarded as fractional ideals, are invertible, then*

$$A = C_1 C_2 \ldots C_n \tag{5}$$

is the unique representation of the integral ideal A as a product of prime ideals.

The proof will be by induction on n. We shall show that for $n = 1$ a representation of the ideal C_1 as a product of two integral ideals different from R,

$$C_1 = B_1 B_2,$$

is not possible at all. For since C_1 is a prime ideal, by § 16.2 we have $B_1 \subseteqq C_1$, say. And since by assumption C_1 is an invertible ideal and the inclusion relation (2) also holds for a fractional ideal A, we have

$$B_2 = RB_2 = C_1^{-1} C_1 B_2 \supseteqq C_1^{-1} B_1 B_2 = C_1^{-1} C_1 = R,$$

which is impossible.

We now proceed to the general case. Suppose that A has the representation (5) as well as the representation

$$A = D_1 D_2 \ldots D_k, \tag{6}$$

where D_j $(j = 1, 2, \ldots, k)$ are prime ideals. Let C_1 be one of the minimal ideals among the C_i $(i = 1, 2, \ldots, n)$. Since A is contained in the prime ideal C_1, by (6) and § 16.2 we have $D_1 \subseteqq C_1$, say. Similarly, there exists an i such that $C_i \subseteqq D_1$. Hence the minimality of C_1 implies that

$$C_i = D_1 = C_1.$$

When we now equate the right-hand sides of (5) and (6) and multiply by C_1^{-1}, we obtain

$$C_2 \ldots C_n = D_2 \ldots D_k;$$

hence, by the induction hypothesis, $n = k$, and after a renumbering, $C_i = D_i$ $(i = 2, \ldots, n)$. //

8. *In a Dedekind ring R every prime ideal is maximal and invertible.*

Let us assume, first of all, that the prime ideal C is invertible and let us show that it is maximal. Suppose that there exists an element $a \in R$ not in C such that the ideal $(C, a) = C + Ra$ is different from R. Then, a fortiori, the ideal $C + Ra^2$ is different from R, because $C + Ra^2 \subseteqq C + Ra$,

and since we are dealing with a Dedekind ring, these two ideals have decompositions into products of prime ideals,

$$C + Ra = \prod_{i=1}^{n} C_i, \qquad C + Ra^2 = \prod_{j=1}^{k} D_j. \tag{7}$$

We now have to construct a ring that in § 18.6 will be called the *factor ring* of R with respect to the ideal C. We decompose the additive group of R into cosets of C. If $C + x$ and $C + y$ are any two cosets, then by multiplying an arbitrary element of the first coset by an arbitrary element of the second we obtain an element of the coset $C + xy$ (because C is an ideal). This enables us to speak of a multiplication of cosets. We can also speak of an addition of cosets, because the sum of an arbitrary element from $C + x$ and an arbitrary element from $C + y$ lies in the coset $C + (x + y)$. It is easy to verify that the set of cosets of C forms a ring $\bar{R}$ under these operations; its zero is C and its unit element is $C + 1$.

In our case, the ring $\bar{R}$ does not have divisors of zero, because from $x \notin C$, $y \notin C$, and

$$(C + x)(C + y) = C$$

it would follow that $xy \in C$, although C is a prime ideal. The results of the preceding paragraph are therefore applicable to $\bar{R}$.

Since by (7) $C \subseteq C_i$ $(i = 1, 2, \ldots, n)$, the prime ideal C_i splits into cosets of C which, as is easy to verify, form a prime ideal $\bar{C}_i$ in $\bar{R}$; similarly, we can introduce prime ideals $\bar{D}_j$ $(j = 1, 2, \ldots, k)$. When we put $\bar{a} = C + a$, then by (7), the principal ideal $\bar{R}\bar{a}$ has the decomposition

$$\bar{R}\bar{a} = \prod_{i=1}^{n} \bar{C}_i$$

and therefore

$$(\bar{R}\bar{a})^2 = \bar{R}\bar{a}^2 = \prod_{i=1}^{n} \bar{C}_i^2 = \prod_{j=1}^{k} \bar{D}_j. \tag{8}$$

All the prime factors that occur in the decomposition (8) of the principal—and hence, by § 16.6, invertible—ideal $\bar{R}\bar{a}^2$ are themselves invertible, so that by § 16.7 the decompositions (8) are identical. Hence it follows that $k = 2n$ and, after a renumbering,

$$C_i = D_{2i-1} = D_{2i} \qquad (i = 1, 2, \ldots, n).$$

Thus,

$$C + Ra^2 = (C + Ra)^2,$$

and hence

$$C \subset (C + Ra)^2 \subset C^2 + Ra.$$

For every $x \in C$, therefore, there exist $y \in C^2$ and $z \in R$ such that

$$x = y + za.$$

Hence $za \in C$, and since C is a prime ideal and $a \notin C$, we have $z \in C$. Thus,

$$C \subseteq C^2 + Ca,$$

and since the converse inclusion is obvious, we have

$$C = C^2 + Ca.$$

Multiplying both sides by the ideal C^{-1} (which exists, because C is invertible), we obtain

$$R = C + Ra,$$

in contradiction to the hypothesis. This proves that C is a maximal ideal.

Now let us show that every prime ideal C is invertible. Let $x \in C$, $x \neq 0$. Then

$$Rx = \prod_{i=1}^{n} C_i,$$

where the ideals C_i $(i = 1, 2, \ldots, n)$ are prime, and, by § 16.6, invertible (because Rx is a principal ideal and is therefore invertible), so that by what has been shown above they are maximal. However, since $Rx \subseteq C$ and C is a prime ideal, we have $C_i \subseteq C$ for at least one i; and since C_i is maximal, $C_i = C$, so that the ideal C is invertible. //

9. From the results of the two preceding sections we derive the following important proposition:

In a Dedekind ring every ideal different from R and O has a unique decomposition into the product of prime ideals.

10. Let us now show that *every non-zero fractional ideal A of a Dedekind ring R splits into the product of positive or negative powers of prime—and hence invertible—ideals.*

For let $A = \frac{1}{d} A_0$. Then

$$A_0 = \prod_{i=1}^{n} C_i, \qquad Rd = \prod_{j=1}^{k} D_j,$$

where all the C_i and D_j are prime; since all the prime ideals D_j ($j = 1, 2, 3, \ldots, k$) are invertible, we have

$$A = A_0(Rd)^{-1} = \prod_{i=1}^{n} C_i \cdot \prod_{j=1}^{k} D_j^{-1}$$

Thus, we have proved the fundamental theorem of § 16.5 in one direction.

11. Now let us prove the converse part of the theorem. If the non-zero fractional ideals of an integral domain R with unit element form a group under multiplication, and are therefore invertible, then, as we know from § 16.6, every (integral) ideal in R is finitely generated. Hence it follows that R *satisfies the maximal condition for ideals*: The union B of an ascending chain of ideals

$$A_1 \subsetneqq A_2 \subsetneqq \ldots \subsetneqq A_n \subsetneqq \ldots$$

is itself an ideal and is therefore generated by a finite number of elements. We can find an n such that all these elements lie in A_n, and then

$$B = A_n = A_{n+1} = \ldots.$$

The statement that R is a Dedekind ring will be proved, in virtue of § 16.2, if we can show that *every integral ideal of R different from R and O is a product of maximal ideals.* If this is not so, then consider those ideals that cannot be so represented and let A be a maximal one of them. A is not a maximal ideal of R, but by the maximal condition there exists a maximal ideal M containing it. Since $A \subsetneqq M$ and $M^{-1}M = R$, we have $M^{-1}A \subsetneqq R$, so that $M^{-1}A$ is an integral ideal. From

$$A = RA = M(M^{-1}A) \tag{9}$$

it follows that

$$A \subsetneqq M^{-1}A.$$

If this were a strict inclusion, then the ideal $M^{-1}A$ would split into the product of maximal ideals, and then by (9), A would also have such a representation, contrary to assumption. Therefore $M^{-1}A = A$, or

$$A = MA. \tag{10}$$

We shall show that (10) leads to a contradiction by showing that there exists an element $z \in M$ for which the product of $1 - z$ by an arbitrary element of A would vanish, although $z \neq 1$ and R has no divisors of zero.

Let A be generated by the elements $x_1, x_2, \ldots, x_n$:

$$A = (x_1, x_2, \ldots, x_n).$$

We put

$$A_i = (x_i, x_{i+1}, \ldots, x_n) \qquad (i = 1, 2, \ldots, n)$$

and $A_{n+1} = O$ and we shall show the existence of a $z_i \in M$ $(i = 1, 2, \ldots, n+1)$ such that

$$(1 - z_i)A \subseteq A_i. \tag{11}$$

Since $A_1 = A$, we can take $z_1 = 0$. Suppose that a z_i with the property (11) has already been found. Then by (10),

$$(1 - z_i)A = (1 - z_i)MA \subseteq MA_i,$$

and since every element of the ideal MA_i can be represented as a sum of products of the elements $x_i, x_{i+1}, \ldots, x_n$ by certain elements of M, we have

$$(1 - z_i)x_i = \sum_{j=i}^{n} z_{ij}x_j \qquad (z_{ij} \in M).$$

Hence

$$(1 - z_i - z_{ii})x_i \in A_{i+1},$$

and therefore

$$(1 - z_{i+1})A \subseteq A_{i+1},$$

where

$$1 - z_{i+1} = (1 - z_i)(1 - z_i - z_{ii}).$$

Clearly, $z_{i+1} \in M$. The element z_{n+1} is then the required element z. This proves the theorem. //

❖ An integral domain R with unit element is a Dedekind ring if and only if it satisfies the following conditions: 1) The maximal condition for ideals; 2) Every prime ideal is maximal; 3) The ring R is *integrally-closed* in its field of fractions P, i.e., every element of P that is a root of a polynomial with coefficients in R and leading coefficient 1 is contained in R. ❖

CHAPTER III

UNIVERSAL ALGEBRAS. GROUPS WITH MULTI-OPERATORS

§ 17. Universal Algebras. Homomorphisms

1. The parallelism between the theory of groups and the theory of rings, which has frequently been in evidence in the preceding chapter, can also be found in many other branches of these theories. In many cases it becomes expedient, instead of discussing groups and rings separately, to construct a unified theory from which results concerning groups and rings can be derived as simple corollaries. It was with this precise purpose that the investigation of algebraic structures with an arbitrary number of algebraic, not necessarily binary, operations was begun.

2. Let G be a set. We shall say that an *n-ary* algebraic operation ω is defined on G (where n is a non-negative integer) if to every ordered system of n elements $a_1, a_2, \ldots, a_n$ of G there is assigned a uniquely determined element of the same set; the result of applying the operation ω to this system of elements will be written as $a_1 a_2 \ldots a_n \omega$. Occasionally it becomes necessary to waive the postulate that the n-ary operation ω should be defined for *every* ordered system of n elements, and the operation is then only *partial*. The postulate that the operation should be single-valued is usually retained, so that the concept of an n-ary operation is only a special case of that of an $(n+1)$-ary relation (see § 7.1).

For $n=2$, we again have our familiar concept of a binary operation (see § 7.2); for $n=3$, we have a *ternary* operation, and so forth. For $n=1$, we shall speak of a *unary operation*. This is an operation that assigns to every element $a \in G$ a uniquely determined element $a\omega \in G$, i.e., it is a single-valued mapping of G into itself. Finally, the case $n=0$, the case of a *nullary operation*, means that in G a certain definite element is fixed that does not depend on the choice of any elements or systems of elements in G. Thus, in choosing the unit element of a group, we are making use of a nullary operation; in the case of a ring, such an operation

would be the choosing of its zero element or of its unit element (provided the latter exists).

3. A set G is called a *universal algebra* if a system Ω of n-ary algebraic operations is given in G, where for distinct operations $\omega \in \Omega$ the numbers n may be distinct or equal. This system of operations may also be infinite: examples of universal algebras of this kind are vector spaces over infinite fields; here we have a single binary operation, namely addition, and an infinite set of unary operations, namely the multiplications by elements of the ground field.

In the preceding chapter we have become acquainted with many distinct types of universal algebras: groupoids, groups, quasigroups, rings, etc. Note that we have two ways of regarding a group as a universal algebra: on the one hand, it is a set with three binary operations, namely multiplication and left and right division; on the other hand, it is a set with a single binary operation, multiplication, a single unary operation, the taking of the inverse of an element, and a single nullary operation, the taking of the unit element. We shall return in § 22.6 to the problem of the various ways of exhibiting one and the same algebraic system as a universal algebra.

Note also that fields can be looked upon as universal algebras only when partial algebraic operations are admitted to consideration, since left and right division in a field and the taking of the inverse of an element are of this kind.

4. Let G be a universal algebra with a system of operations Ω. A subset $A \subseteq G$ is called a *subalgebra* of G if for every n-ary operation $\omega \in \Omega$ it always follows from $a_1, a_2, \ldots, a_n \in A$ that

$$a_1 a_2 \ldots a_n \omega \in A.$$

Special cases of this concept are, obviously, a subgroupoid of a groupoid, a subgroup of a group, and a subring of a ring. We note that if a ring R has a unit element and if R is regarded as a universal algebra whose operations include the nullary operation of taking the unit element, then subalgebras are only those subrings of R (see § 9.2) that contain the unit element of R, and not arbitrary subrings, even if they possess unit elements of their own.

Just as in § 9.6, it can be shown that *the intersection of an arbitrary system of subalgebras of a universal algebra G is a subalgebra, provided it is not empty.*

From this it follows that *if M is an arbitrary non-empty subset of a universal algebra, then there exists a uniquely determined subalgebra $\{M\}$ that is minimal among the subalgebras wholly containing M.* This is the

intersection of all subalgebras of G containing M, one of these subalgebras being G itself. If $\{M\} = G$, then M is a *system of generators* for G.

5. Two universal algebras G and G' in which systems of operations Ω and Ω', respectively, are given, are said to be *of the same type* if a one-to-one correspondence between Ω and Ω' can be set up under which every operation $\omega \in \Omega$ and the operation $\omega' \in \Omega'$ corresponding to it are n-ary with one and the same n. We may therefore assume that universal algebras of the same type have one and the same system of operations.

Universal algebras G and G' of the same type with one and the same system of operations Ω are called *isomorphic* if there exists a one-to-one mapping φ of G onto G' such that for every n-ary operation $\omega \in \Omega$ and arbitrary elements $a_1, a_2, \ldots, a_n \in G$, we have

$$(a_1 a_2 \ldots a_n \omega)\varphi = (a_1\varphi)(a_2\varphi) \ldots (a_n\varphi)\omega. \tag{1}$$

The concept of isomorphism plays exactly the same role for universal algebras as it did for groups, rings, and partially ordered sets. We shall now introduce a generalization of this concept which is very important for the whole subsequent theory, namely the concept of *homomorphism*. When we retain the condition (1), but only assume that φ is a single-valued (and not necessarily one-to-one) mapping of G into G' (i.e., not necessarily onto the whole of the latter) we arrive at the concept of a *homomorphic mapping* of one universal algebra into another of the same type.

From the definition of a homomorphism it follows easily that *the product of homomorphisms* (in the sense of § 1.2) *is itself a homomorphism*. For if G, G', G'' are algebras of the same type with one and the same system Ω of operations and if $\varphi\colon G \to G'$ and $\psi\colon G' \to G''$ are homomorphisms, then for every n-ary operation $\omega \in \Omega$ and arbitrary elements $a_1, a_2, \ldots, a_n \in G$, we have

$$\begin{aligned}(a_1 a_2 \ldots a_n \omega)(\varphi\psi) &= [(a_1 a_2 \ldots a_n \omega)\varphi]\psi = \\ &= [(a_1\varphi)(a_2\varphi) \ldots (a_n\varphi)\omega]\psi = [(a_1\varphi)\psi]\,[(a_2\varphi)\psi] \ldots [(a_n\varphi)\psi]\omega = \\ &= [a_1(\varphi\psi)]\,[a_2(\varphi\psi)] \ldots [a_n(\varphi\psi)]\omega,\end{aligned}$$

which proves that the product $\varphi\psi$ is the homomorphism of G into G''. //

From (1) it also follows that if $G\varphi$ is the image of G under the homomorphic mapping φ into G', then $G\varphi$ *is a subalgebra of* G'.

If $G\varphi = G'$, then we speak of a homomorphic mapping *onto* G' and we call G' a *homomorphic image* of G.

6. Let us apply the concept of homomorphism to the case of groupoids and rings. In the case of a binary multiplication the equation (1), of course, becomes

$$(ab)\varphi = a\varphi \cdot b\varphi, \tag{2}$$

and a universal algebra of the same type as a groupoid is itself a groupoid. It is easy to verify, as in the case of isomorphism (see § 10.1), that in a homomorphic mapping of a groupoid G onto a groupoid G' such properties of the operation in G as commutativity and associativity are preserved.

Let φ be a homomorphic mapping of a groupoid G onto a groupoid G'. If G has a unit element e, then $e\varphi$ is the unit element in G'. Also, if the element $b \in G$ is one of the (right) inverse elements of $a \in G$, then $b\varphi$ is one of the (right) inverse elements of a.

In fact, for $a \in G$ it follows from $ae = ea = a$ and (2) that $a\varphi \cdot e\varphi = e\varphi \cdot a\varphi = a\varphi$, and since $a\varphi$ ranges over the whole of G' when a ranges over G, $e\varphi$ is the unit element of G'. Furthermore, it follows from $ab = e$ and (2) that $a\varphi \cdot b\varphi = e\varphi$. //

Thus, *the homomorphic image of a semigroup and of a group is a semigroup and a group, respectively.*

Let us now examine a homomorphic mapping φ of a ring R onto a universal algebra R' of the same type. From what has been said above, it follows that R' is an abelian group under addition and a groupoid under multiplication. Let us show that the distributive laws hold in R'. If a', b', $c' \in R'$, and if the elements a, b, $c \in R$ are such that $a\varphi = a'$, $b\varphi = b'$, $c\varphi = c'$, then the equations

$$(a + b)c = ac + bc,$$

which hold in R, imply that

$$(a' + b')c' = a'c' + b'c'.$$

The second distributive law is verified in the same way.

Thus, *the homomorphic image of a ring is a ring.* From what has been shown above it follows that under a homomorphism of rings the zero goes over into the zero, and that the homomorphic image of an associative ring and of a commutative ring is a ring having the same property.

7. There is a method of surveying all homomorphic mappings of a given universal algebra G. Let us first introduce the following concepts.

We consider a universal algebra G with a system of operations Ω. An equivalence relation π (see § 3.1) given on G is called a *congruence* on G if for an arbitrary n-ary operation $\omega \in \Omega$ and arbitrary elements

$$a_i, a_i' \in G \qquad (i = 1, 2, \ldots, n)$$

it follows from

$$a_i \pi a_i' \qquad (i = 1, 2, \ldots, n)$$

that

$$(a_1 a_2 \ldots a_n \omega) \pi (a_1' a_2' \ldots a_n' \omega).$$

In other words, if $A_1, A_2, \ldots, A_n$ are arbitrary classes of the partition defined by the equivalence relation π (see § 3.2), then the class B containing the element $a_1 a_2 \ldots a_n \omega$ ($a_i \in A_i$, $i = 1, 2, \ldots, n$) does not depend on how a_i is chosen in A_i ($i = 1, 2, \ldots, n$). This enables us to define an n-ary operation ω in the factor set G/π (see § 3.4) by putting

$$A_1 A_2 \ldots A_n \omega = B. \tag{3}$$

Since this holds for all operations $\omega \in \Omega$, this converts G/π into a universal algebra with the same system of operations Ω as the original algebra G. This algebra G/π is called a *factor algebra* of G with respect to π.

The definition (3) of an operation in the factor algebra G/π and the definition of a homomorphism (1) show that the natural mapping of an algebra G onto a factor algebra G/π (see § 1.4) is a homomorphism. It is called the *natural homomorphism* of G onto G/π.

From the existence of a natural homomorphism of G onto G/π and what has been said in the preceding paragraph, it follows that *factor algebras of groupoids, groups, and rings are themselves groupoids, groups, and rings, respectively.* We can therefore speak of the *factor groupoid* (or *factor group*, or *factor ring*) G/π of a groupoid (group, ring) G with respect to a congruence π.

8. The converse to what was shown in the preceding subsection is the following *homomorphism theorem*, which gives a survey of all homomorphic mappings of universal algebras.

If G is a universal algebra with a system of operations Ω and φ is a homomorphic mapping of G onto a universal algebra G' of the same type, $G\varphi = G'$, then there exists a congruence π on G such that the algebra G' is isomorphic to the factor algebra G/π. Moreover, there exists an isomorphic mapping ψ of G' onto G/π such that the product $\varphi\psi$ coincides with the natural homomorphism of G onto G/π.

For we obtain a partition of G into disjoint classes if we refer to a single class those elements of G whose images under φ coincide. The equivalence relation π defined by this partition is a congruence on G. In fact, for every n-ary operation $\omega \in \Omega$ it follows from

$$a_i \varphi = \bar{a}_i \varphi \qquad (i = 1, 2, \ldots, n)$$

by (1) that

$$(a_1a_2\dots a_n\omega)\varphi = (a_1\varphi)(a_2\varphi)\dots(a_n\varphi)\omega = \\ = (\bar{a}_1\varphi)(\bar{a}_2\varphi)\dots(\bar{a}_n\varphi)\omega = (\bar{a}_1\bar{a}_2\dots\bar{a}_n\omega)\varphi,$$

i.e., the elements $a_1a_2\dots a_n\omega$ and $\bar{a}_1\bar{a}_2\dots\bar{a}_n\omega$ also belong to a single class of the partition π.

The factor algebra G/π therefore exists. By assigning to every element $a'\in G'$ the class A of π formed from all the inverse images of the element a' under the homomorphism φ, we obtain a one-to-one mapping ψ of G' onto G/π. Let us show that ψ is an isomorphism.

Let ω be an arbitrary n-ary operation of Ω and a_i' $(i=1, 2, \dots, n)$, arbitrary elements of G'. When we put

$$a_i'\psi = A_i \qquad (i=1, 2, \dots, n)$$

and choose elements $a_i\in A_i$ $(i=1, 2, \dots, n)$ for which

$$a_i\varphi = a_i' \qquad (i=1, 2, \dots, n),$$

then, by (1), we obtain

$$(a_1a_2\dots a_n\omega)\varphi = a_1'a_2'\dots a_n'\omega. \tag{4}$$

Since by (3),

$$a_1a_2\dots a_n\omega\in A_1A_2\dots A_n\omega,$$

it follows from (4) that

$$(a_1'a_2'\dots a_n'\omega)\psi = A_1A_2\dots A_n\omega = (a_1'\psi)(a_2'\psi)\dots(a_n'\psi)\omega,$$

which shows that ψ is an isomorphism.

Finally, if a is an arbitrary element of G and if

$$a\varphi = a', \qquad a'\psi = A,$$

then $a\in A$, by the definition of ψ. This shows that the product $\varphi\psi$ coincides with the natural homomorphism of G onto G/π. //

§ 18. Groups with Multi-operators

1. As we shall see below, there exists a very close connection between congruences, and hence homomorphisms, of groups and rings on the one hand, and normal subgroups of groups and ideals of rings, on the other. This connection cannot be carried over in its fullest extent to arbitrary universal algebras nor, in particular, to groupoids or semigroups. How-

ever, it is entirely preserved for one special class of universal algebras that was recently introduced by P. J. Higgins (*Proc. London Math. Soc.*, vol. 6 (1956), pp. 366-416) and represents a very successful unification of the classes of groups and rings.

Let G be a group. It need not be commutative, but we shall find it convenient to write it in *additive notation*; in particular, the null element of the group will be denoted, as usual, by the symbol 0. G is called a *group with a system of multi-operators* Ω—or, for brevity, an *Ω-group*—if apart from addition a certain system of n-ary algebraic operations Ω is given on G (for several n with $n \geqq 1$), provided that for all $\omega \in \Omega$ we have

$$00 \ldots 0\omega = 0, \tag{1}$$

where the element 0 on the left is repeated n times if ω is an n-ary operation.

Clearly, for an empty system of operations we are back to the concept of a group. Furthermore, the concept of an Ω-group turns into that of a ring if the *additive group* of the Ω-group is commutative and if the system of operations Ω consists of a single binary multiplication connected with addition by the distributive laws; as we already know, the condition (1) then holds.

2. An Ω-group G can be regarded as a universal algebra under the operations of the additive group and those of Ω. Every subalgebra of G (see § 17.4) is itself an Ω-group, because as a subgroup of the additive group it contains 0 and condition (1) continues to hold. We shall therefore speak not of subalgebras but of *Ω-subgroups* of an Ω-group G.

From this it follows by § 17.4 that *the intersection of an arbitrary system of Ω-subgroups of an Ω-group G is itself an Ω-subgroup* and, in particular, contains the element 0. An Ω-subgroup is also the subalgebra $\{M\}$ generated by a non-empty subset M of G.

We note that by (1) the zero subgroup of the additive group is an Ω-subgroup.

3. If an Ω-group G is mapped by a homomorphism φ (see § 17.5) onto a universal algebra G' of the same type, then this gives, in particular, a homomorphism of the additive group of G, so that G', under the operations corresponding to those of the additive group of G, is itself a group. We shall write this group additively and denote its zero by $0'$. Since φ is a homomorphism and $0\varphi = 0'$, we have for every operation $\omega \in \Omega$, by (1),

$$0'0' \ldots 0'\omega = 0'.$$

Thus, *every homomorphic image of an Ω-group is itself an Ω-group.* In particular, *factor algebras of Ω-groups* (see § 17.7) *are themselves Ω-groups.*

We may therefore speak of the Ω-factor-group G/π of an Ω-group G with respect to a congruence π.

4. A non-empty subset A of an Ω-group G is called an *ideal* in G if the following two conditions are satisfied:

1) A is a normal subgroup of the additive group;

2) When $\omega \in \Omega$ is an arbitrary n-ary operation, a an arbitrary element of A, and $x_1, x_2, \ldots, x_n$ arbitrary elements of G, then the following inclusion relation must hold for $i = 1, 2, \ldots, n$:

$$-(x_1 x_2 \ldots x_n \omega) + x_1 \ldots x_{i-1}(a + x_i)x_{i+1} \ldots x_n \omega \in A. \tag{2}$$

For groups, the concept of an ideal defined just above coincides with that of a normal subgroup, because condition (2) is void for an empty system of operations Ω.

For rings, our new concept of an ideal coincides with that of a (two-sided) ideal introduced in § 13.8.

For in this case condition 1) requires that A should be a subgroup of the additive group of the ring, and for the operation of multiplication and for arbitrary $a \in A$, $x_1, x_2, \in G$, condition 2) becomes, for $i = 1$:

$$-x_1 x_2 + (a + x_1)x_2 = -x_1 x_2 + a x_2 + x_1 x_2 = a x_2 \in A,$$

and for $i = 2$:

$$-x_1 x_2 + x_1(a + x_2) = -x_1 x_2 + x_1 a + x_1 x_2 = x_1 a \in A,$$

as required. //

Note that the inclusion relation (2) can be rewritten in the form

$$x_1 \ldots x_{i-1}(a + x_i)x_{i+1} \ldots x_n \omega \in x_1 x_2 \ldots x_n \omega + A \quad (i = 1, 2, \ldots, n), \tag{3}$$

where the element on the right-hand side is the coset of the normal subgroup A generated by the element $x_1 x_2 \ldots x_n \omega$. Applying (3) several times, we arrive at the following proposition:

For every ideal A of an Ω group G, every n-ary operation $\omega \in \Omega$, arbitrary elements $a_1, a_2, \ldots, a_n \in A$, and arbitrary elements $x_1, x_2, \ldots, x_n \in G$ we have

$$(a_1 + x_1)(a_2 + x_2) \ldots (a_n + x_n)\omega \in x_1 x_2 \ldots x_n \omega + A. \tag{4}$$

Hence it follows that *every ideal A of an Ω-group G is an Ω-subgroup of G.* For A is a subgroup of the additive group, by condition 1); and

for $x_1 = x_2 = \ldots = x_n = 0$ it follows from (4) for every n-ary operation $\omega \in \Omega$ and arbitrary $a_1, a_2, \ldots, a_n \in A$, on account of (1), that

$$a_1 a_2 \ldots a_n \omega \in A.$$

Clearly, the ideals of an Ω-group G include, in particular, G itself and the zero subgroup 0. If G has no other ideals, then it is called a *simple* Ω-group.

It is not difficult to verify that *the intersection of an arbitrary system of ideals of an Ω-group G is itself an ideal,* so that we can speak of the ideal *generated* by an arbitrary system of elements M.

Note that *the ideal generated by a system of ideals A_i $(i \in I)$ of an Ω-group G coincides with the subgroup B of the additive group generated by these ideals.*

For we know from § 13.5 that $B = \{A_i, i \in I\}$ is a normal subgroup of the additive group and that every element of B can be written in the form

$$b = a_1 + a_2 \ldots + a_k \qquad (a_j \in A_{i_j},\ j = 1, 2, \ldots, k). \tag{5}$$

Suppose now that an n-ary operation $\omega \in \Omega$, an element $b \in B$ with the expression (5), and elements $x_1, x_2, \ldots, x_n \in G$ are given. Since $A_{i_1}, A_{i_2}, A_{i_3}, \ldots, A_{i_k}$ are ideals, we obtain with the help of (3) and (5)

$$x_1 \ldots x_{i-1}(b + x_i)x_{i+1} \ldots x_n \omega \in x_1 x_2 \ldots x_n \omega + B,$$

as required. //

The ideal $B = \{A_i, i \in I\}$ is called the *sum* of the ideals A_i $(i \in I)$.

5. We shall presently prove a theorem that illustrates the role of ideals in the theory of Ω-groups. Note that when we speak of the *decomposition of an Ω-group G with respect to an ideal A,* we have in mind the decomposition of the additive group of the Ω-group with respect to A as a normal subgroup.

There is a one-to-one correspondence between the decompositions of an arbitrary Ω-group G with respect to its ideals and the congruences on G.

For let A be an arbitrary ideal of an Ω-group G. By the definition of a normal subgroup, there exists, for any $a_1, a_2 \in A$ and $x_1, x_2 \in G$, an element $a_3 \in A$ such that

$$a_1 + x_2 = x_2 + a_3,$$

and therefore

$$(x_1 + a_1) + (x_2 + a_2) = (x_1 + x_2) + (a_3 + a_2) \in (x_1 + x_2) + A.$$

On the other hand, for every n-ary operation $\omega \in \Omega$ and for arbitrary elements $a_1, a_2, \ldots, a_n \in A$ and $x_1, x_2, \ldots, x_n \in G$ the inclusion (4) holds. This shows that the decomposition of G into cosets of A is, in fact, a congruence on the Ω-group G.

Now let π be an arbitrary congruence on G. Let A denote that class of the partition π containing the zero element of the additive group; the elements $a \in A$ are therefore characterized by the fact that $a \pi 0$ holds. If $a_1, a_2 \in A$, then $a_1 \pi 0$ and $a_2 \pi 0$; and therefore, by the definition of a congruence,

$$(a_1 + a_2)\pi(0 + 0), \qquad \text{or} \qquad (a_1 + a_2)\pi 0;$$

and hence $a_1 + a_2 \in A$. Next, if $a \in A$, then

$$[0 + (-a)]\pi[a + (-a)],$$

or $-a \pi 0$, and hence $-a \in A$. Finally, if $a \in A$, $x \in G$, then

$$(-x + a + x)\pi(-x + 0 + x),$$

or $(-x + a + x)\pi 0$; and hence $-x + a + x \in A$. This proves that A is a normal subgroup of the additive group.

Now if an n-ary operation $\omega \in \Omega$, an element $a \in A$, and elements $x_1, x_2, x_3, \ldots, x_n \in G$ are given, then

$$(a + x_i)\pi(0 + x_i), \qquad \text{or} \qquad (a + x_i)\pi x_i;$$

hence

$$[x_1 \ldots x_{i-1}(a + x_i)x_{i+1} \ldots x_n \omega]\pi(x_1 x_2 \ldots x_n \omega),$$

and therefore

$$-x_1 x_2 \ldots x_n \omega + x_1 \ldots x_{i-1}(a + x_i)x_{i+1} \ldots x_n \omega \in A \qquad (i = 1, 2, \ldots, n).$$

Thus we see that the class A is even an ideal of the Ω-group G.

Finally, let us examine an arbitrary class B of the partition π. If $b \in B$, $a \in A$, i.e. $a \pi 0$, then

$$(b + a)\pi(b + 0), \qquad \text{or} \qquad (b + a)\pi b,$$

which implies for the coset $b + A$ the inclusion

$$b + A \subseteq B.$$

On the other hand, if b' is an arbitrary element of B, then from $b' \pi b$ it follows that $(-b + b')\pi 0$, or

$$-b + b' \in A; \qquad \text{i.e.,} \qquad b' \in b + A.$$

This shows that $B = b + A$ and proves that every class of the partition π is a coset of the ideal A. //

6. In virtue of this theorem, from now on we shall speak not of the Ω-factor-group G/π of an Ω-group G with respect to a congruence π, but of the Ω-factor-group with respect to an ideal A and we shall denote it by G/A. The ideal A is, of course, the zero element of this Ω-factor-group.

By applying all that has been said to the case of groups, we see that *there is a one-to-one correspondence between the congruences on a group G and the decompositions with respect to its various normal subgroups.* We can therefore speak of the *factor group of a normal subgroup A in a group G*, which we shall denote by G/A. Similarly, *there is a one-to-one correspondence between the congruences in a ring R and the decompositions with respect to its various (two-sided) ideals,* so that we can speak of the *factor ring of an ideal A in a ring R*, which we shall denote by R/A.

One further consequence of the theorem proved above is the following. If φ is a homomorphic mapping of an Ω-group G onto an Ω-group G', then the *kernel* of φ is defined as the collection of those elements of G that are mapped under φ into the zero element of G'. From the theorem of § 18.5 and the Homomorphism Theorem of § 17.8, we now obtain the following theorem.

There is a one-to-one correspondence between the kernels of the homomorphisms of an Ω-group and the ideals of that Ω-group. A homomorphic image of an Ω-group is determined to within isomorphism by the kernel of the homomorphism in question.

7. Let us examine a few very simple, but important examples.

Every homomorphic image of a cyclic group is itself cyclic.

For if G' is the image of a cyclic group $G = \{a\}$ under a homomorphism φ and $a\varphi = a'$, then for every element $g' \in G'$ we can find an inverse image $a^k \in G$, so that

$$g' = a^k\varphi = (a\varphi)^k = (a')^k;$$

or, in other words, $G' = \{a'\}$. //

Now let us find all the homomorphic images of an infinite cyclic group, which we can choose, by § 10.2, as the additive group of integers. In § 10.2 we have given a survey of the subgroups of this group, all of which, of course, are normal. If $\{n\}$ is the subgroup formed by the numbers that are multiples of the natural number n, then two numbers belong to the same coset of this subgroup if and only if their difference is divisible by n, i.e., if the numbers give one and the same remainder on division by n. Since the remainders on division of integers by n can only be the numbers 0, 1,

$2, \ldots, n-1$, the factor group of the additive group of integers with respect to the subgroup $\{n\}$ is a finite group of order n and, as was shown above, is cyclic. Now n was an arbitrary number, so that when we take into account the description of cyclic groups given in § 10.2, we arrive at the following result:

The homomorphic images of an infinite cyclic group are the totality of all possible cyclic groups.

8. Now let us find all the homomorphic images of the ring of integers C. As mentioned in § 13.8, all the subgroups of the additive group of this ring are ideals of the ring, so that it remains for us to examine the factor rings of these ideals.

The factor ring of the zero ideal is, of course, isomorphic to C itself. Let C_n denote the factor ring of C with respect to the ideal formed by the multiples of the natural number n, which we call *the residue-class ring modulo n*. From what has been shown in the preceding subsection, it follows that this ring is finite and consists of n elements, that its additive group is cyclic, and that we can choose the numbers $0, 1, 2, \ldots, n-1$ as representatives of the cosets forming the ring. Correspondingly, the elements of C_n can be denoted by $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}$.

From the definition of the operations in a factor ring (see § 17.7) we derive the following rule for operating in C_n: if $0 \leqq k, l < n$ and $k + l = nq_1 + r_1$, $kl = nq_2 + r_2$, where $0 \leqq r_1, r_2 < n$, then

$$\bar{k} + \bar{l} = \bar{r}_1, \qquad \bar{k} \cdot \bar{l} = \bar{r}_2.$$

In particular, $\bar{0}$ is the zero element of C_n and $\bar{1}$ the unit element.

The ring C_n has divisors of zero when n is a composite number, but is a commutative field when n is a prime number.

For if $n = kl$ $(1 < k, l < n)$, then the elements $\bar{k}$ and $\bar{l}$ of C_n are different from $\bar{0}$, but $\bar{k} \cdot \bar{l} = \bar{0}$. If n is a prime number, $n = p \geqq 2$, then every number k satisfying the inequality $1 \leqq k \leqq p-1$ is prime to p. Therefore, there exist integers l and m such that

$$kl + pm = 1,$$

and l can be chosen so that $1 \leqq l \leqq p-1$. Hence it follows that in the residue-class ring C_p the equation

$$\bar{k} \cdot \bar{l} = \bar{1}$$

holds, i.e., every element of C_p, other than the zero element, has an inverse in C_p, so that this ring is a commutative field. //

The residue-class rings C_p with respect to a prime modulus p ($p = 2, 3, 5, 7, 11, 13, \ldots$) are our first examples of finite fields, and they play a very important role in the theory of fields, as will be shown presently.

9. Let K be a field, not necessarily associative (see § 12.1). The intersection of all subfields of K is itself a subfield, which we call provisionally the *prime field* of K. Our object is a description of the structure of this subfield.

We define the *center* of an arbitrary ring R as the set of all elements a of R that are permutable with every element of R, that is,

$$ax = xa \tag{6}$$

for all $x \in R$, and in addition, for all $x, y \in R$, satisfy the conditions

$$(ax)y = a(xy), \qquad (xy)a = x(ya) \tag{7}$$

and hence also

$$(xa)y = x(ay),$$

because by (6) and (7)

$$(xa)y = (ax)y = a(xy) = (xy)a = x(ya) = x(ay).$$

The zero element of R and also its unit element, provided it exists, obviously belong to the center of the ring.

The center of a ring R is an associative-commutative subring. The center of a field is a commutative subfield.

For if two elements a and b belong to the center of a ring R, then for arbitrary $x, y \in R$,

$$(a \pm b)x = ax \pm bx = xa \pm xb = x(a \pm b),$$

$$[(a \pm b)x]y = (ax \pm bx)y = (ax)y \pm (bx)y = \\ = a(xy) \pm b(xy) = (a \pm b)(xy);$$

the second condition (7) can be verified in the same way. Furthermore,

$$(ab)x = a(bx) = a(xb) = (ax)b = (xa)b = x(ab),$$

$$[(ab)x]y = [a(bx)]y = a[(bx)y] = a[b(xy)] = (ab)(xy).$$

Thus, the center of R proves to be a subring, which is obviously associative-commutative.

If R is a field with the unit element e and if a is an element of its center, with $a \neq 0$, then by solving the equations $ax = e$ and $ya = e$ and bearing

in mind that a belongs to the center, we find a uniquely determined left and right inverse element a^{-1}. Then for arbitrary $x, y \in R$,

$$a^{-1}x = (a^{-1}x)(aa^{-1}) = [(a^{-1}x)a]a^{-1} = [a^{-1}(xa)]a^{-1} =$$
$$= [a^{-1}(ax)]a^{-1} = [(a^{-1}a)x]a^{-1} = xa^{-1},$$

$$(a^{-1}x)y = (a^{-1}a)[(a^{-1}x)y] = a^{-1}\{a[(a^{-1}x)y]\} =$$
$$= a^{-1}\{[a(a^{-1}x)]y\} = a^{-1}\{[(aa^{-1})x]y\} = a^{-1}(xy).$$

The second condition (7) can be verified in the same way. This shows that a^{-1} belongs to the center of R, so that the center is a commutative field.

10. Generally speaking, the center of a ring may happen to consist of the zero element only. If K is a field with unit element e, then its center necessarily contains e and also all the multiples ne, where n is an arbitrary integer. All these multiples form an associative-commutative subring $C^{(K)}$ of K that is contained in all the subfields of K and therefore in the prime field of K. On the other hand, the prime field of K must be contained in the center of K, because the center is a commutative subfield, and so the *prime field of K is associative and commutative.*

The prime field of K is isomorphic either to the field of rational numbers or to one of the residue-class fields C_p with respect to a prime modulus p.

To prove this, we consider the mapping $n \to ne$ of the ring of integers C onto $C^{(K)}$. This mapping is a homomorphism, so that, by what was shown in § 18.8, the ring $C^{(K)}$ is isomorphic either to the ring of integers C or to some residue-class ring C_n—in fact, inasmuch as there are no divisors of zero in K, even to a residue-class field C_p with respect to a prime modulus p. In the latter case, the field C_p is then the required prime field of K. But in the first case, the center of K, which is a commutative field and contains the subring $C^{(K)}$ isomorphic to the ring of integers, also contains the field of fractions of this subring, which is isomorphic (see § 11.7) to the field of rational numbers; this is then the prime field of K. //

11. If the prime field of K is isomorphic to the field of rational numbers, then we say that K is *without characteristic* (or that it has *characteristic zero*); all number fields, in particular, are of this kind. But if the prime field of K is isomorphic to the residue-class field C_p with respect to a prime modulus p, then K is called a field of *finite characteristic*, or *prime characteristic*, and specifically, *of characteristic p.*

Every non-zero element of a field K without characteristic (of characteristic p) has infinite order (order p) in the additive group of the field.

For if K is without characteristic and $a \in K$ $(a \neq 0)$, then an equation $na = 0$ would lead to $(ne)a = 0$, and hence $ne = 0$—because K has no

divisors of zero—and so $n = 0$. But if the characteristic of K is p, then for every a of K we have

$$pa = (pe)a = 0 \cdot a = 0,$$

so that for $a \neq 0$ the order of this element is equal to the prime number p. //

12. We now return to the discussion of an arbitrary Ω-group G. We choose an ideal A in G and consider its Ω-factor-group $G' = G/A$. Let B' be an arbitrary Ω-subgroup of the Ω-group G' and B its *complete inverse image* in G under the natural homomorphism of G onto G', i.e., the set of those elements of G contained in the cosets constituting B'. Let us show that *B is an Ω-subgroup of G.*

For if $x, y \in B$, i.e., $x + A, y + A \in B'$, then

$$(x + y) + A = (x + A) + (y + A) \in B',$$
$$-x + A = -(x + A) \in B',$$

i.e., $x + y \in B$, $-x \in B$. On the other hand, for an n-ary operation $\omega \in \Omega$ and elements $x_1, x_2, \ldots, x_n \in B$, the relations $x_i + A \in B'$, the equation $x_i + A = A + x_i$ $(i = 1, 2, \ldots, n)$, and (4) yield

$$x_1 x_2 \ldots x_n \omega + A = (x_1 + A)(x_2 + A) \ldots (x_n + A)\omega \in B',$$

and hence $x_1 x_2 \ldots x_n \omega \in B$. //

The Ω-subgroup B just constructed obviously contains the ideal A. Conversely, *if B is an arbitrary Ω-subgroup of an Ω-group G containing the ideal A, then its image B' under the natural homomorphism of G onto $G' = G/A$ is an Ω-subgroup and B is the complete inverse image of B'.*

For the natural homomorphism of G onto G' induces a homomorphism of B into G' and therefore, by what was proved in § 17.5, B' is an Ω-subgroup of G'. On the other hand, it follows from $B \supseteq A$ that every element of G that occurs in a coset of A belonging to B' is contained in B. //

Thus we have proved the following theorem.

By assigning to every Ω-subgroup of $G' = G/A$ its complete inverse image under the natural homomorphism of G onto G', we obtain a one-to-one correspondence, preserving the inclusion relation, between all the Ω-subgroups of G' and all those Ω-subgroups of G that contain A.

13. We supplement this result by the following theorem.

Under the correspondence indicated in the preceding theorem, an ideal of one of the Ω-groups G, G' corresponds to an ideal of the other, and the Ω-factor-groups with respect to corresponding ideals are isomorphic.

For if B' is an ideal of the Ω-group $G'=G/A$ and B its complete inverse image in G, then—as follows from what was shown in § 17.5—the result of carrying out the natural homomorphisms of G onto G' and of G' onto $G''=G'/B'$ is a homomorphism of G onto G''. The kernel of this homomorphism (see § 18.6) is constituted by those elements of G that are mapped into B' under the mapping of G onto G', that is, the elements that form the Ω-subgroup B. Hence it follows, by § 18.6, § 18.5, and § 17.8, that B is an ideal of G and that the Ω-factor-groups G/B and G'/B' are isomorphic.

Suppose now that B is an ideal of an Ω-group G and B' its image under the natural homomorphism of G onto G'. If $b\in B$, $x\in G$, then $b+A\in B'$, $x+A\in G'$. Since

$$-x+b+x\in B,$$

we have

$$-(x+A)+(b+A)+(x+A)=(-x+b+x)+A\in B',$$

i.e., B' is a normal subgroup of the additive group of G'. But if an n-ary operation $\omega\in\Omega$ and arbitrary elements $b\in B$ and $x_1, x_2, \ldots, x_n\in G$ are given —that is, $b+A\in B'$, $x_i+A\in G'$ $(i=1, 2, \ldots, n)$—then from

$$-x_1x_2\ldots x_n\omega+x_1\ldots x_{i-1}(b+x_i)x_{i+1}\ldots x_n\omega=b_0\in B$$
$$(i=1, 2, \ldots, n)$$

it follows that

$$-(x_1+A)(x_2+A)\ldots(x_n+A)\omega+$$
$$+(x_1+A)\ldots(x_{i-1}+A)[(b+A)+(x_i+A)](x_{i+1}+A)\ldots(x_n+A)\omega=$$
$$=b_0+A\in B' \qquad (i=1, 2, \ldots, n).$$

This proves that B' is an ideal in G'. //

Note that in the last part of the proof we have not made use of the assumption that $B\supseteq A$.

§ 19. Automorphisms and Endomorphisms. The Field of p-adic Numbers

1. Let G be a universal algebra. Every isomorphic mapping of G onto itself is called an *automorphism*. Thus, G always has the *identity automorphism*: the identity mapping of G onto itself. Examples of non-trivial automorphisms are the automorphism of the additive group of integers that carries every integer k into $-k$ and the automorphism of the field of complex numbers that carries every complex number $a+bi$ into $a-bi$.

If several automorphisms of an algebra G are carried out in succession, the result is itself an automorphism. Under this multiplication, which by § 1.2 is associative, all the automorphisms of G form a group, because the identity automorphism is the unit element and the inverse mapping of an arbitrary automorphism is also an automorphism; this group is called the *automorphism group* of the universal algebra G.

❖ Every group is isomorphic to the group of all automorphisms of a suitable universal algebra [G. Birkhoff, *Revista Unione Mat. Argentina*, vol. 11 (1946), pp. 155-57].

Every group can be embedded isomorphically in the group of automorphisms of a suitable abelian group. ❖

2. Let G be a non-commutative semigroup with unit element and let ε be a *divisor of unity* in G, that is, an element that has a left and right inverse ε^{-1} in G (see § 14.1). When we *transform* G by ε, i.e., when we map every element x of G into $\varepsilon^{-1}x\varepsilon$, we obtain an automorphism of G, a so-called *inner automorphism*. For it follows from

$$\varepsilon^{-1}x\varepsilon = \varepsilon^{-1}y\varepsilon$$

that $x = y$, that is, the mapping in question is one to one. Further, from

$$x = \varepsilon^{-1}(\varepsilon x\varepsilon^{-1})\varepsilon$$

it follows that this is a mapping onto the whole of G. Finally, the equation

$$\varepsilon^{-1}(xy)\varepsilon = \varepsilon^{-1}x\varepsilon \cdot \varepsilon^{-1}y\varepsilon$$

shows that the mapping is an isomorphism.

Clearly, a commutative semigroup with unit element—in particular, every abelian group—has only one inner automorphism, namely the identity.

It is easy to verify that the product of the transformation by ε of a semigroup with unit element and the transformation by δ coincides with the transformation by $\varepsilon\delta$ and that the transformation by ε^{-1} is inverse to the transformation by ε. Hence it follows that *the inner automorphisms of a semigroup with unit element form a subgroup in the group of all automorphisms of the semigroup.*

The subgroup of inner automorphisms is even normal in the group of all automorphisms of the semigroup G in question. For let φ be an arbitrary automorphism of G and α an inner automorphism generated by the divisor of unity ε. Then, for every $x \in G$,

$$x(\varphi^{-1}\alpha\varphi) = [\varepsilon^{-1}(x\varphi^{-1})\varepsilon]\varphi = (\varepsilon^{-1})\varphi \cdot (x\varphi^{-1})\varphi \cdot (\varepsilon\varphi) = (\varepsilon\varphi)^{-1}x\varepsilon\varphi,$$

so that the automorphism $\varphi^{-1}a\varphi$ is an inner automorphism and is generated by the divisor of unity $\varepsilon\varphi$. //

Moreover, if G is a group, so that all the elements of G are divisors of unity, then we obtain a homomorphic mapping of the whole group G onto the group of its inner automorphisms if we assign to every element of G the inner automorphism generated by it. The kernel of this homomorphism (see § 18.6) is the set of elements of G that are permutable with every element of the group; it is called the *center* of G. From the homomorphism theorem it follows that *the center of a group G is a normal subgroup and the group of inner automorphisms of G is isomorphic to the factor group of G with respect to its center.*

3. The concept of an inner automorphism extends to the case of an arbitrary associative ring with unit element: *Every inner automorphism of the multiplicative semigroup of an associative ring R with unit element is an automorphism of the ring itself,* because

$$\varepsilon^{-1}(x+y)\varepsilon = \varepsilon^{-1}x\varepsilon + \varepsilon^{-1}y\varepsilon.$$

Such an automorphism is naturally called an *inner automorphism* of the ring R.

4. Every homomorphic mapping of a universal algebra G into itself is called an *endomorphism* of G. The endomorphisms include, in particular, all the automorphisms and also all the isomorphic mappings of G into itself and all the homomorphic mappings of G onto itself.

The result of carrying out several endomorphisms in succession is itself an endomorphism, so that all the endomorphisms form a semigroup under this multiplication, which by § 1.2 is associative; this is the so-called *semigroup of endomorphisms* of the universal algebra G.

This semigroup of endomorphisms has a unit element: the identity automorphism. The divisors of unity of the semigroup of endomorphisms are the automorphisms; and they constitute the only divisors of unity, because a single-valued inverse mapping is only possible in the case of automorphisms.

It is worth noting that the identity automorphism of a universal algebra G, in addition to playing the role of a unit element for the endomorphisms of the algebra, is obviously also a *left unit* for all the homomorphic mappings of G into every other algebra and a *right unit* for all the homomorphic mappings of other algebras into G.

5. Let us now continue our discussion of homomorphic mappings of Ω-groups. If φ and ψ are two homomorphisms of an Ω-group G into an Ω-group G', then the mapping

$$a(\varphi+\psi) = a\varphi + a\psi \qquad (a\in G) \tag{1}$$

is not, in general, a homomorphism. It is a homomorphism (see § 17.5) if and only if for arbitrary $a, b \in G$

$$(a+b)(\varphi+\psi) = a(\varphi+\psi) + b(\varphi+\psi),$$

i.e.,

$$b\varphi + a\psi = a\psi + b\varphi, \tag{2}$$

and if, for an arbitrary n-ary operation $\omega \in \Omega$ and arbitrary $a_1, a_2, \ldots, a_n \in G$, we have

$$(a_1 a_2 \ldots a_n \omega)(\varphi+\psi) = [a_1(\varphi+\psi)]\,[a_2(\varphi+\psi)] \ldots [a_n(\varphi+\psi)]\,\omega,$$

i.e.,

$$(a_1\varphi)(a_2\varphi)\ldots(a_n\varphi)\omega + (a_1\psi)(a_2\psi)\ldots(a_n\psi)\omega = \\ = (a_1\varphi + a_1\psi)(a_2\varphi + a_2\psi)\ldots(a_n\varphi + a_n\psi)\omega. \tag{3}$$

When the conditions (2) and (3) are satisfied, we call the homomorphisms φ and ψ *summable* and the homomorphism $\varphi+\psi$, their *sum*.

In the case of groups without multi-operators, the condition (2) shows that *homomorphisms φ and ψ of a group G into a group G' are summable if and only if the subgroups $G\varphi$ and $G\psi$ of G' are elementwise permutable.* If homomorphisms φ and ψ of a ring R into a ring R' are considered, then condition (2) is automatically satisfied and (3) becomes the following condition: for arbitrary $a, b \in R$

$$a\varphi \cdot b\varphi + a\psi \cdot b\psi = (a\varphi + a\psi)(b\varphi + b\psi),$$

i.e.,

$$a\varphi \cdot b\psi + a\psi \cdot b\varphi = 0.$$

6. We have defined a *partial* addition of homomorphisms of an Ω-group G into an Ω-group G'. This addition is commutative, because it follows from (2) that for all $a \in G$,

$$a\varphi + a\psi = a\psi + a\varphi,$$

or, by (1), $\varphi+\psi = \psi+\varphi$. It is also associative, because for all $a \in G$

$$a[(\varphi+\psi)+\chi] = a[(\varphi+(\psi+\chi)] = a\varphi + a\psi + a\chi.$$

Hence it follows that if the sum $\sum_{i=1}^{n} \varphi_i$ of homomorphisms $\varphi_i : G \to G'$ $(i = 1, 2, \ldots, n)$ is a homomorphism, then it is uniquely determined. Note that we can speak of the sum of an *infinite set of homomorphisms* $\varphi_i : G \to G'$ $(i \in I)$ if we assume in addition that for every $a \in G$ only a finite number of the elements $a\varphi_i$ $(i \in I)$ are different from zero.

If the homomorphisms φ and ψ of an Ω-group G into an Ω-group G' are summable, then for arbitrary Ω-groups H and F and arbitrary homomorphisms $\sigma\colon H \to G$ and $\tau\colon G' \to F$ the homomorphisms $\sigma\varphi$ and $\sigma\psi$, and also $\varphi\tau$ and $\psi\tau$ (see § 17.5), *are summable and*

$$\sigma(\varphi + \psi) = \sigma\varphi + \sigma\psi, \qquad (\varphi + \psi)\tau = \varphi\tau + \psi\tau. \tag{4}$$

For if $a \in H$, then

$$a[\sigma(\varphi + \psi)] = (a\sigma)(\varphi + \psi) = (a\sigma)\varphi + (a\sigma)\psi = \\ = a(\sigma\varphi) + a(\sigma\psi) = a(\sigma\varphi + \sigma\psi),$$

that is, the mapping $\sigma\varphi + \sigma\psi$ is a homomorphism and the first equation of (4) holds. It is just as easy to prove the second statement of the theorem. Note that the existence of the sums on the right-hand sides of (4) does not imply the existence of the sums on the left-hand sides. //

7. In virtue of the equation $0 + 0 = 0$, where 0 is the zero element of an Ω-group, and of the equation (1) of § 18.1, the mapping that carries every element of an Ω-group G into the zero element of an Ω-group G' is a homomorphism; this is the *zero homomorphism* of G into G'.

The following statements are readily verified: Every homomorphism $\varphi\colon G \to G'$ is summable with the zero homomorphism and the sum is equal to φ; if the homomorphisms $\varphi, \psi\colon G \to G'$ are summable and

$$\varphi + \psi = \varphi,$$

then ψ is the zero homomorphism; finally, for arbitrary Ω-groups H and F and arbitrary homomorphisms $\sigma\colon H \to G$ and $\tau\colon G' \to F$, the product of σ with the zero homomorphism of G into G' is equal to the zero homomorphism of H into G', and the product of the zero homomorphism of G into G' with τ is equal to the zero homomorphism of G into F.

8. Let us examine the homomorphisms of a group G (without multi-operators) into an *abelian* group G'; we shall write the first group multiplicatively and the second, additively. In this case, by (2), any two homomorphisms are summable. On the other hand, for every homomorphism $\varphi\colon G \to G'$, the mapping $-\varphi$ defined by the equation

$$a(-\varphi) = -a\varphi \qquad (a \in G)$$

is in this case a homomorphism, because for arbitrary $a, b \in G$,

$$(ab)(-\varphi) = -(ab)\varphi = -(a\varphi + b\varphi) = (-a\varphi) + (-b\varphi) = \\ = a(-\varphi) + b(-\varphi).$$

This homomorphism is the opposite of φ, because for $a \in G$

$$a[\varphi + (-\varphi)] = a\varphi + a(-\varphi) = a\varphi - a\varphi = 0,$$

i.e., $\varphi + (-\varphi)$ is equal to the zero homomorphism.

Thus, *the homomorphisms of an arbitrary group G into an abelian group G' form an abelian group under addition*.

Let us discuss, in particular, the endomorphisms of an abelian group G. They form an abelian group under addition and a semigroup with unit element under multiplication (see § 19.4), and the distributive laws (4) are satisfied. Thus, we see that *under the operations of addition and multiplication of endomorphisms, the endomorphisms of an abelian group G form an associative ring with unit element.* This ring is called *the endomorphism ring* of the abelian group G.

Every associative ring can be isomorphically embedded in the endomorphism ring of a suitable abelian group.

By § 10.4, we may assume that the associative ring R under discussion has a unit element 1. If a is an arbitrary element of R, then the mapping that carries every element x of R into xa is an endomorphism of the additive group of R, because

$$(x + y)a = xa + ya.$$

The sum and the product of elements of R correspond to the sum and the product of the corresponding endomorphisms, as the equations

$$x(a + b) = xa + xb,$$
$$x(ab) = (xa)b$$

show. Finally, distinct elements of R correspond to distinct endomorphisms, because it follows from $a \neq b$ that $1 \cdot a \neq 1 \cdot b$. //

9. *The endomorphism ring of an infinite cyclic group is isomorphic to the ring of integers C.*

For if a is a generating element of the given group, written additively, and if the endomorphism φ carries a into ka, where k is an integer, then

$$(na)\varphi = nka, \tag{5}$$

so that the endomorphism φ is completely determined by the number k. Conversely, the mapping φ of our group defined by (5) is, in fact, an endomorphism. Thus, we have established a one-to-one correspondence between the endomorphisms of the infinite cyclic group $\{a\}$ and the integers. That this is an isomorphism follows from the fact that if

$$a\varphi = ka, \qquad a\psi = la,$$

then

$$a(\varphi + \psi) = a\varphi + a\psi = ka + la = (k + l)a,$$

$$a(\varphi\psi) = (a\varphi)\psi = (ka)\psi = (kl)a. \text{ //}$$

Since the only numbers in the ring of integers that have inverses are 1 and -1, *the automorphism group of the infinite cyclic group* $\{a\}$ *is a cyclic group of order two* and consists of the identity automorphism and the automorphism that carries every element of the group into its opposite.

⁂ The ring of endomorphisms of a finite cyclic group of order n is isomorphic to the residue-class ring C_n modulo n.

The endomorphism ring of the additive group of rational numbers is isomorphic to the field of rational numbers. Thus, every non-zero endomorphism of this group is an automorphism. ⁂

10. A very important role in the theory of abelian groups is played by the *group of type* p^∞, where p is a prime number: this is the multiplicative group of all complex p^n-th roots of unity ($n = 1, 2, \ldots$). It is well known that the group of roots of unity of degree p^n, where n is fixed, is cyclic of order p^n with a generating element

$$g_n = \cos(2\pi/p^n) + i \sin(2\pi/p^n).$$

A group of type p^∞ is therefore abelian and infinite; for it is the union of the ascending sequence of cyclic groups $\{g_n\}$ ($n = 1, 2, \ldots$). The elements

$$g_n \qquad (n = 1, 2, \ldots) \tag{6}$$

form a system of generators of this group. If we make use of additive, instead of multiplicative notation for a group of type p^∞, then the generators (6) are related by the equations

$$pg_1 = 0, \qquad pg_{n+1} = g_n \qquad (n = 1, 2, \ldots). \tag{7}$$

Let us find the endomorphism ring of a group of type p^∞. If φ is an endomorphism of this group, then it is completely determined by the images of all the generating elements (6). Since g_n is of order p^n, the order of the element $g_n\varphi$ cannot exceed p^n. However, all these elements lie in the cyclic subgroup $\{g_n\}$, so that

$$g_n\varphi = k_n g_n \qquad (n = 1, 2, \ldots), \tag{8}$$

where

$$0 \leqq k_n < p^n. \tag{9}$$

Further, the equations (7), which hold for the elements (6), must also be satisfied for their images under φ, so that

$$p(g_{n+1}\varphi) = g_n\varphi,$$

and hence

$$p(k_{n+1}g_{n+1}) = k_{n+1}g_n = k_ng_n.$$

From this it follows that the difference $k_{n+1} - k_n$ must be divisible by the order p^n of g_n or, in the notation of the theory of numbers,

$$k_{n+1} \equiv k_n \pmod{p^n}. \tag{10}$$

Thus, to every endomorphism φ of a group of type p^∞ there corresponds a sequence of non-negative integers

$$(k_1, k_2, \ldots, k_n, \ldots), \tag{11}$$

subject to the conditions (9) and (10). Here, two distinct endomorphisms correspond to distinct sequences of the form (11), because at least one of the generators (6) has distinct images under these endomorphisms.

Conversely, every sequence of the form (11) subject to the conditions (9) and (10), corresponds to an endomorphism φ of a group of type p^∞. For on the basis of (8) we can define a mapping φ for all elements of our group, and it is easy to verify that this is an endomorphism.

If in a group of type p^∞ two endomorphisms, φ and ψ, are given by the sequences (11) and

$$(l_1, l_2, \ldots, l_n, \ldots), \tag{12}$$

respectively, then

$$g_n(\varphi + \psi) = (k_n + l_n)g_n,$$
$$g_n(\varphi\psi) = (k_nl_n)g_n.$$

Since the conditions of type (10) hold for the sequences (11) and (12), it follows that

$$k_{n+1} + l_{n+1} \equiv k_n + l_n \pmod{p^n},$$
$$k_{n+1}l_{n+1} \equiv k_nl_n \pmod{p^n};$$

these congruences remain valid when their left-hand sides are replaced by positive residues mod p^{n+1} (i.e., by the remainders on division by p^{n+1}) and the right-hand sides, similarly, by residues mod p^n, and the conditions (9) are then also satisfied.

Thus, the ring of endomorphism of a group of type p^∞ is isomorphic to the ring of sequences of non-negative integers of the form (11) subject to the conditions (9) and (10), where addition and multiplication of these

sequences is performed componentwise, with a subsequent replacement of the n-th term by its residue mod p^n. This commutative ring is called *the ring of p-adic integers*. Its zero is the sequence $(0, 0, \ldots, 0, \ldots)$, corresponding to the zero endomorphism of the group of type p^∞, and its unit element is the sequence $(1, 1, \ldots, 1, \ldots)$, corresponding to the identity automorphism of the group.

11. Another notation for the p-adic integers is often used. Let α be a p-adic integer given by the sequence of non-negative integers (11), subject to the conditions (9) and (10). If we set

$$a_0 = k_1, \qquad a_n = (k_{n+1} - k_n)/p^n \qquad (n = 1, 2, \ldots), \tag{13}$$

then all the a_n $(n = 0, 1, 2, \ldots)$ are integers; in fact they are positive residues modulo p, i.e.,

$$0 \leqq a_n < p \qquad (n = 0, 1, 2, \ldots). \tag{14}$$

Since by (13),

$$k_n = a_0 + a_1 p + a_2 p^2 + \ldots + a_{n-1} p^{n-1} \qquad (n = 1, 2, \ldots), \tag{15}$$

the p-adic integer α can be associated with the infinite series

$$a_0 + a_1 p + a_2 p^2 + \ldots + a_n p^n + \ldots ; \tag{16}$$

distinct p-adic numbers obviously correspond to distinct series of the form (16). Conversely, if the coefficients of an arbitrary series of the form (16) are integers subject to (14), then, when we define k_n by the equations (15), we obtain a sequence of the form (11) subject to the conditions (9) and (10).

Thus, we have set up a one-to-one correspondence between all the p-adic integers and all the series of the form (16) whose coefficients are residues modulo p. Using (15) we can easily carry over the operations defined fo. p-adic integers to series of the form (16): if α is given by (16) and β by

$$b_0 + b_1 p + b_2 p^2 + \ldots + b_n p^n + \ldots ,$$

also subject to the conditions (14), then

$$\alpha + \beta = c_0 + c_1 p + c_2 p^2 + \ldots + c_n p^n + \ldots ,$$

where all the coefficients c_n are residues modulo p and

$$c_0 = a_0 + b_0 - p q_0,$$
$$c_n = a_n + b_n + q_{n-1} - p q_n \quad (n = 1, 2, \ldots);$$

moreover,

$$\alpha\beta = d_0 + d_1 p + d_2 p^2 + \ldots + d_n p^n + \ldots ,$$

where the coefficients d_n are also residues modulo p and

$$d_0 = a_0 b_0 - p s_0,$$

$$d_n = \sum_{k+l=n} a_k b_l + s_{n-1} - p s_n \qquad (n = 1, 2, \ldots).$$

12. From this description of the multiplication it follows immediately, since p is a prime number, that *the ring of p-adic integers does not contain divisors of zero,* so that, by § 11.5, it has a field of fractions, the so-called *field of p-adic numbers.* This field can be constructed as follows.

We consider all possible series of the form

$$a_k p^k + a_{k+1} p^{k+1} + \ldots + a_n p^n + \ldots, \tag{17}$$

where all the coefficients a_n are residues modulo p and k is greater than, equal to, or less than, zero; the series (17) can therefore contain only a finite number of terms with negative powers of p. If not all the coefficients of (17) are zero, then we shall take $a_k \neq 0$; series with zero coefficients are assumed to be identical.

When the definition of the operations on series of the form (16) is carried over in the natural way to series of the form (17), we see that the latter series form an associative-commutative ring whose unit element is the series of the form (17) in which $a_0 = 1$ and all the remaining coefficients are zero. This ring is even a field: if the series (17) is different from zero, i.e., $a_k \neq 0$, then its inverse is the series

$$b_{-k} p^{-k} + b_{-k+1} p^{-k+1} + \ldots + b_n p^n + \ldots,$$

whose coefficients, being residues modulo p, can be found successively from the following equations, which are solvable, because p is a prime number:

$$\begin{aligned}
a_k b_{-k} - p s_{-k} &= 1, \\
a_k b_{-k+1} + a_{k+1} b_{-k} + s_{-k} - p s_{-k+1} &= 0, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_k b_n + a_{k+1} b_{n-1} + \ldots + a_{n+2k} b_{-k} + s_{n-1} - p s_n &= 0, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

The series of the form (17) in which $k \geqq 0$ form a subring in our field that is isomorphic to the ring of p-adic integers. And the whole field is the field of fractions of this subring: every series of the form (17) after multiplication (in the sense of our definition of this operation) by a suitable positive power of p (which is, of course, a p-adic integer) becomes a series of the form (16), that is, a p-adic integer.

§ 20. Normal Series and Composition Series

1. *If A is an ideal and B is an Ω-subgroup of an Ω-group G, then the Ω-subgroup $\{A, B\}$ generated by A and B consists of those elements of G that can be expressed in at least one way in the form $a+b$, where $a \in A$, and $b \in B$; hence it is meaningful to write*

$$\{A, B\} = A + B. \tag{1}$$

For it is clear that every element of the form $a+b$ belongs to $\{A, B\}$. Let us show that these elements form an Ω-subgroup containing, of course, A and B. Indeed, if $a, a' \in A$, $b, b' \in B$, then $b + a' - b \in A$, because A is a normal subgroup of the additive group. Therefore

$$(a+b) + (a'+b') = [a + (b + a' - b)] + (b + b') = a'' + b'',$$

where $a'' \in A$, $b'' \in B$. Further, $0 = 0 + 0$, but the zero element is contained both in A and in B. And the opposite element of $a+b$ is $a'-b$, where $a' = -b - a + b \in A$. Finally, for every n-ary operation $\omega \in \Omega$ and arbitrary elements

$$a_1, a_2, \ldots, a_n \in A \quad \text{and} \quad b_1, b_2, \ldots, b_n \in B,$$

we have from (4) of § 18.4 that

$$(a_1 + b_1)(a_2 + b_2) \ldots (a_n + b_n)\omega \in b_1 b_2 \ldots b_n \omega + A;$$

and since

$$b_1 b_2 \ldots b_n \omega = b \in B,$$

B being an Ω-group, $b + A = A + b$. //

2. The following *isomorphism theorem* is used very frequently.

If A and B are Ω-subgroups of an Ω-group G and A is an ideal in the Ω-subgroup $\{A, B\}$, then the intersection $A \cap B$ is an ideal in B, and we have the following isomorphism:

$$\{A, B\}/A \cong B/(A \cap B). \tag{2}$$

For the equation (1) shows that every coset of A in $\{A, B\}$ contains at least one element of B. Thus, in the natural homomorphism of $\{A, B\}$ onto $\{A, B\}/A$, the Ω-subgroup B is mapped homomorphically onto the whole of this Ω-factor-group. The kernel of the homomorphism (see § 18.6) is obviously $A \cap B$. It is, therefore, an ideal in B and the validity of the isomorphism (2) follows from the theorems of § 17.8 and § 18.5. //

3. We shall now prove *Zassenhaus' Lemma,* a generalization of the isomorphism theorem that will be used in an essential way later in the proof of Schreier's Theorem.

If A, A', B, and B' are Ω-subgroups of an Ω-group G, where A' and B' are ideals in A and B, respectively, then $A'+(A\cap B')$ and $B'+(B\cap A')$ are ideals, respectively, in $A'+(A\cap B)$ and $B'+(B\cap A)$ and we have the isomorphism

$$A'+(A\cap B)/A'+(A\cap B')\cong B'+(B\cap A)/B'+(B\cap A'). \tag{3}$$

When $A\supseteq B$ and $B'=O$, this lemma obviously becomes the Isomorphism Theorem.

To prove the lemma, we set

$$C=A\cap B.$$

Since B' is an ideal in B and $C\subseteq B$, by the Isomorphism Theorem

$$C\cap B'=A\cap B\cap B'=A\cap B'$$

is an ideal in C. The same is true for the intersection $B\cap A'$ and therefore for the sum

$$C'=(A\cap B')+(B\cap A') \tag{4}$$

of these two ideals (see § 18.4). We put

$$D=C/C'.$$

Furthermore, since A' is an ideal in A, we have by § 20.1,

$$\{A', A\cap B\}=A'+(A\cap B)=A'+C.$$

An arbitrary element of this sum is of the form $a'+c$, where $a'\in A'$, $c\in C$. Let us associate it with the coset (i.e., element of D) $C'+c$. If the element $a'+c$ has another expression of the same form,

$$a'+c=a_1'+c_1 \qquad (a_1'\in A',\ c_1\in C),$$

then

$$-a_1'+a'=c_1-c\in A'\cap C\subseteq A'\cap B\subseteq C',$$

and therefore

$$c_1=(-a_1'+a')+c\in C'+c.$$

Thus we obtain a single-valued mapping of the Ω-group $A'+C$ into the Ω-group D, and in fact onto the whole of this Ω-group, because every element $c\in C$, which obviously belongs to $A'+C$, is mapped into the coset

$C'+c$. This mapping is a homomorphism: since A' is an ideal in $A'+C$, we have

$$(a_1'+c_1)+(a_2'+c_2)=a_3'+(c_1+c_2) \qquad (a_3'\in A')$$

and, by § 18.4 (4), for every n-ary operation $\omega\in\Omega$

$$(a_1'+c_1)(a_2'+c_2)\dots(a_n'+c_n)\omega=a_0'+c_1c_2\dots c_n\omega,$$

where $a_0'\in A'$ and $c_1c_2\dots c_n\omega\in C$, because C is an Ω-group.

The kernel of the homomorphism so constructed is the sum $A'+(A\cap B')$. For this sum belongs to the kernel, since $A\cap B'\subseteq C'$. On the other hand, if an element $a'+c\in A'+C$ is mapped under the homomorphism in question into C', then $c\in C'$ and therefore, by (4), we can write:

$$c=u+v, \qquad \text{where} \qquad u\in B\cap A', \quad v\in A\cap B';$$

clearly, the terms in the expression of a sum of two ideals can be written down in any order. Thus,

$$a'+c=(a'+u)+v\in A'+(A\cap B').$$

Hence it follows that $A'+(A\cap B')$ is an ideal in $A'+C=A'+(A\cap B)$ and furthermore that we have the isomorphism

$$A'+(A\cap B)/A'+(A\cap B')\cong D.$$

By a symmetry argument we can state that $B'+(B\cap A')$ is an ideal in $B'+(B\cap A)$ and that

$$B'+(B\cap A)/B'+(B\cap A')\cong D. \;/\!/$$

4. A finite sequence of Ω-subgroups of an Ω-group G each embedded in the following,

$$G=A_0\supset A_1\supset A_2\supset\dots\supset A_k=O, \tag{5}$$

beginning with G itself and ending with the zero element is called a *normal series* of the group if every A_i $(i=1, 2, \dots, k)$ is a proper ideal of A_{i-1} (but not necessarily of A_j for $j<i-1$). The number k is called the *length* of the normal series (5), and the Ω-factor-groups

$$G/A_1,\ A_1/A_2,\ \dots,\ A_{k-1}/A_k=A_{k-1}$$

are called the *factors* of the series.

Every Ω-group G has normal series: for example, $G\supset O$ and, provided G has a non-trivial ideal A, $G\supset A\supset O$.

A normal series

$$G = B_0 \supset B_1 \supset B_2 \supset \ldots \supset B_l = O$$

is called a *refinement* of (5) if every A_i $(i = 1, 2, \ldots, k-1)$ coincides with one of the Ω-subgroups B_j; clearly, $l \geqq k$.

Two normal series of an Ω-group G are called *isomorphic* if their lengths are equal and if a one-to-one correspondence can be set up between their factors such that corresponding factors are isomorphic Ω-groups. It is not assumed that this correspondence preserves the disposition of the factors, i.e., the i-th factor of the first series does not have to be isomorphic to the i-th factor of the second series.

5. SCHREIER'S THEOREM: *Any two normal series of an arbitrary Ω-group G have isomorphic refinements.*

For suppose that G has the two normal series

$$G = A_0 \supset A_1 \supset A_2 \supset \ldots \supset A_k = O, \tag{6}$$

$$G = B_0 \supset B_1 \supset B_2 \supset \ldots \supset B_l = O. \tag{7}$$

We set

$$A_{ij} = A_i + (A_{i-1} \cap B_j) \qquad (i = 1, 2, \ldots, k;\ j = 0, 1, \ldots l);$$

$$B_{ji} = B_j + (B_{j-1} \cap A_i) \qquad (j = 1, 2, \ldots, l;\ i = 0, 1, \ldots k).$$

These expressions have a meaning on account of § 20.1, because, for example, A_i is an ideal and $A_{i-1} \cap B_j$ is an Ω-subgroup of A_{i-1}. Note that for $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, l$ we have the inclusions

$$A_{i-1} = A_{i0} \supseteq A_{i,j-1} \supseteq A_{ij} \supseteq A_{il} = A_i,$$

$$B_{j-1} = B_{j0} \supseteq B_{j,i-1} \supseteq B_{ji} \supseteq B_{jk} = B_j.$$

By Zassenhaus' Lemma (see § 20.3), A_{ij} and B_{ji} are ideals in $A_{i,j-1}$ and $B_{j,i-1}$, respectively, and the corresponding Ω-factor-groups are isomorphic:

$$A_{i,j-1}/A_{ij} \cong B_{j,i-1}/B_{ji}. \tag{8}$$

If all the A_{ij} $(j = 1, 2, \ldots, l-1)$ are inserted between A_{i-1} and A_i $(i = 1, 2, \ldots, k)$ in (6), then we obtain, in general, a refinement of (6) *with repetitions*, because it may happen that $A_{i,j-1} = A_{ij}$. Similarly, we can obtain a refinement of (7) with repetitions. By (8), these refinements are isomorphic.

To conclude the proof, it remains to go over to refinements without repetitions. If $A_{i,j-1} = A_{ij}$, i.e.,

$$A_{i,j-1}/A_{ij} = O,$$

then by (8) we also have $B_{j,i-1} = B_{jt}$. Hence it follows that without disturbing the isomorphism of the refinements in question we may simultaneously delete all repetitions from them. //

❖ An *ascending normal series* of an Ω-group G is defined as a system of Ω-subgroups A_α satisfying the following conditions:

1. The indices α form a well-ordered set (see § 5.4) with the first element 0 and the last element μ.
2. $A_0 = O$, $A_\mu = G$.
3. If $\alpha < \beta$, then $A_\alpha \subset A_\beta$.
4. If we denote the index immediately following α by $\alpha + 1$, then A_α is an ideal in $A_{\alpha+1}$ for all α.
5. If α is a limit number, then A_α is the set-theoretical union of all the A_β for which $\beta < \alpha$.

Any two ascending normal series of an arbitrary Ω-group G have isomorphic refinements that are also ascending normal series. [A. G. Kurosh, *Mat. Sb.*, vol. 16 (1945), pp. 59-72]. The example of the infinite cyclic group shows that for infinite *descending* normal series the analogous theorem does not hold. ❖

6. A normal series of an Ω-group that has no refinements other than the series itself is called a *composition series* of the group. By § 18.13, *a normal series of an Ω-group is a composition series if and only if all the factors are simple Ω-groups* (see § 18.4).

Schreier's Theorem has the following two corollaries:

THE JORDAN-HÖLDER THEOREM: *If an Ω-group G has composition series, then any two composition series are isomorphic.*

If an Ω-group G has composition series, then each of its normal series can be refined to a composition series and therefore has a length not exceeding the length of the composition series of the group.

To prove the second theorem it is sufficient to apply Schreier's Theorem to the given normal series and to one of the composition series of the group in question.

When we apply the above concepts to the case of groups without multi-operators, we observe that all finite groups as well as some infinite groups have composition series. However, neither the infinite cyclic group nor a group of type p^∞ (see § 19.10) have composition series.

7. An *invariant series* of an Ω-group G is a finite system of ideals of G ordered by inclusion, beginning with G and ending with the zero element. Invariant series are a special case of normal series, and therefore the concept of an isomorphism of series and of refinements of a series (see § 20.4)

have a meaning for it. An invariant series without proper refinements is called a *principal series.* We have the following theorems:

Any two invariant series of an arbitrary Ω-group can be refined to isomorphic invariant series.

If an Ω-group has principal series, then any two of its principal series are isomorphic and every invariant series can be refined to a principal series.

It is easy to see that these theorems can be proved by the same methods as were used for the corresponding theorems for normal series. But they can also be derived immediately from the theorems proved above if we make use of the following construction. Let G be an Ω-group. We extend the system of multi-operators Ω to a system Ω', by adding to Ω a certain, in general infinite, system of unary operations. Namely, we add to Ω all the inner automorphisms of the additive group of G and all the mappings of G into itself defined as follows: for every n-ary operation $\omega \in \Omega$, arbitrary elements $x_1, x_2, \ldots, x_n \in G$ and arbitrary i $(1 \leqq i \leqq n)$, we take the mapping that carries every element $a \in G$ into

$$-x_1 x_2 \ldots x_n \omega + x_1 \ldots x_{i-1}(a + x_i) x_{i+1} \ldots x_n \omega.$$

All the additional unary operations carry zero into zero, and so satisfy the condition (1) of § 18.1, and hence G is an Ω'-group. By the definition of an ideal (see § 18.4) *the Ω'-subgroups of G are the only ideals of G,* so that normal (composition) series of G as an Ω'-group coincide with the invariant (principal) series of G as an Ω-group. //

§ 21. Abelian, Nilpotent, and Solvable Ω-groups

1. Let A and B be Ω-subgroups of an Ω-group G. The *commutator group* $[A, B]$ of these Ω-subgroups is defined as the ideal of the Ω-subgroup $\{A, B\}$ generated by the set of all elements of the following two forms:

$$[a, b] = = -a - b + a + b \qquad (a \in A,\ b \in B) \tag{1}$$

—this element is called the *commutator* of a and b—and

$$[a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \omega] = $$
$$= = -a_1 a_2 \ldots a_n \omega - b_1 b_2 \ldots b_n \omega + (a_1 + b_1)(a_2 + b_2) \ldots (a_n + b_n)\omega, \tag{2}$$

where ω is an n-ary operation of Ω, $a_1, a_2, \ldots, a_n \in A$, and $b_1, b_2, \ldots, b_n \in B$.

Thus, in the case of a group without multi-operators, the commutator group $[A, B]$ of two subgroups $A, B \subseteq G$ is the normal subgroup of $\{A, B\}$ generated by all possible commutators $[a, b]$, where $a \in A$, $b \in B$. If we con-

sider a ring R, then $[a, b]$ is always equal to zero and the elements (2) assume the form

$$[a_1, a_2; b_1, b_2] = -a_1a_2 - b_1b_2 + (a_1 + b_1)(a_2 + b_2) =$$
$$= a_1b_2 + b_1a_2 = [a_1, 0; 0, b_2] + [0, a_2; b_1, 0].$$

In this case, therefore, the commutator $[A, B]$ of two subrings $A, B \subseteq R$ is the ideal of the subring $\{A, B\}$ generated by all possible products ab and ba, where $a \in A$, $b \in B$.

From the definition of an ideal it follows immediately that if A is an Ω-subgroup of an Ω-group G and M is a subset of G, then the ideal generated by M in A is contained in the ideal generated by M in G. Hence it follows that *if A', B', A, B are Ω-subgroups of an Ω-group G and $A' \subseteq A$, $B' \subseteq B$, then*

$$[A', B'] \subseteq [A, B].$$

❖ In a group without multi-operators, the commutator group $[A, B]$ coincides with the subgroup generated by all commutators $[a, b]$ $(a \in A,\ b \in B)$. ❖

2. *For arbitrary Ω-subgroups A and B of an Ω-group G, we have the equality*

$$[A, B] = [B, A]. \tag{3}$$

Indeed, for arbitrary $b \in B$, $a \in A$,

$$[b, a] = -a + [-a, b] + a \in [A, B], \tag{4}$$

because $[A, B]$ is a normal subgroup of $\{A, B\}$ under addition. Moreover, for every n-ary operation $\omega \in \Omega$ and for arbitrary $b_1, b_2, \ldots, b_n \in B$, $a_1, a_2, \ldots, a_n \in A$, we have

$$b_1b_2 \ldots b_n\omega \in B, \qquad a_1a_2 \ldots a_n\omega \in A;$$

and therefore, by (4),

$$[b_1b_2 \ldots b_n\omega,\ a_1a_2 \ldots a_n\omega] \in [A, B].$$

Further,

$$b_i + a_i = [-b_i, -a_i] + a_i + b_i \qquad (i = 1, 2, \ldots, n),$$

and since, by (4),

$$[-b_i, -a_i] \in [A, B] \qquad (i = 1, 2, \ldots, n)$$

and $[A, B]$ is an ideal in $\{A, B\}$, we have by (4) of § 18.4,

$$(b_1 + a_1)(b_2 + a_2) \dots (b_n + a_n)\omega = (a_1 + b_1)(a_2 + b_2) \dots (a_n + b_n)\omega + d,$$

where $d \in [A, B]$. Thus we see that

$$[b_1, b_2, \dots, b_n; a_1 a_2, \dots, a_n; \omega] =$$
$$= [b_1 b_2 \dots b_n \omega, a_1 a_2 \dots a_n \omega] +$$
$$+ [a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_n; \omega] + d \in [A, B]. \tag{5}$$

(4) and (5) imply the inclusion

$$[B, A] \subseteq [A, B],$$

and (3) follows from this by a symmetry argument. //

3. As the definition shows, the commutator group $[A, G]$ of an arbitrary Ω-subgroup A with the Ω-group G is an ideal in G. This is true, in particular, for

$$G' = [G, G].$$

This ideal is called the *derived group* of the Ω-group G.

For a group G without multi-operators, the derived group G' is therefore the normal subgroup generated by all commutators $[a, b]$ $(a, b \in G)$. For a ring R, the corresponding formation is the ideal generated by all products ab, where $a, b \in R$; in the theory of rings this ideal is usually called the *square* of R.

An Ω-subgroup A of an Ω-group G is an ideal of G if and only if

$$[A, G] \subseteq A. \tag{6}$$

For if A is an ideal in G, then for all $a \in A$, $x \in G$ the element $-x + a + x$, and therefore also the element $[a, x]$ is contained in A. Furthermore, by (4) of § 18.4, for every n-ary operation $\omega \in \Omega$ and all $a_1, a_2, \dots, a_n \in A$, $x_1, x_2, \dots, x_n \in G$, the element

$$-x_1 x_2 \dots x_n \omega + (a_1 + x_1)(a_2 + x_2) \dots (a_n + x_n)\omega$$

and hence also the element

$$[a_1, a_2, \dots, a_n; x_1, x_2, \dots, x_n; \omega]$$

is contained in A. Consequently, the ideal A also contains the ideal $[A, G]$ generated by all these elements.

Conversely, if (6) is satisfied, then A contains, in particular, all the commutators $[a, x]$ $(a \in A, x \in G)$, and hence $-x + a + x \in A$, so that condi-

tion 1) of § 18.4 is satisfied. Condition 2) is also satisfied: for, (2) of § 18.4 follows from

$$[0, \ldots, 0, a, 0, \ldots, 0; x_1, x_2, \ldots, x_n; \omega] \in A,$$

because $0 \ldots 0a0 \ldots 0\omega \in A$. //

From this theorem it follows, in virtue of the inclusion relation

$$[A, G] \subseteqq [G, G] = G',$$

that *every Ω-subgroup A of an Ω-group G containing the derived group G' is an ideal in G.*

4. An Ω-group G is called *abelian* if its derived group is zero:

$$[G, G] = O.$$

This means, in particular, that for all $a, b \in G$

$$[a, b] = 0,$$

from which it follows that

$$a + b = b + a, \tag{7}$$

i.e., *an abelian Ω-group has an abelian additive group.* On the other hand, for every n-ary operation $\omega \in \Omega$ and arbitrary $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in G$, we have

$$[a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \omega] = 0,$$

so that, by (7),

$$(a_1 + b_1)(a_2 + b_2) \ldots (a_n + b_n)\omega = a_1 a_2 \ldots a_n \omega + b_1 b_2 \ldots b_n \omega. \tag{8}$$

For groups without multi-operators this concept becomes that of an ordinary abelian group; and for rings it becomes that of a ring with zero multiplication, because for rings the condition (8) is equivalent to $ab = 0$ for all a and b.

Every Ω-subgroup A of an abelian Ω-group G is an ideal in G, because $A \supseteqq G' = O$.

Every Ω-subgroup A and every Ω-factor-group G/A of an abelian Ω-group G are themselves abelian, because the conditions (7) and (8), being valid in G, are also valid in A and G/A.

5. *An Ω-factor-group G/A of an Ω-group G is abelian if and only if the ideal A contains the derived group G' of G:*

$$A \supseteqq G'. \tag{9}$$

For the fact that G/A is abelian means that for all $x, y \in G$

$$[x + A, y + A] = A$$

and for every n-ary operation $\omega \in \Omega$ and arbitrary $x_1, x_2, \ldots, x_n, y_1, y_2, y_3, \ldots, y_n \in G$

$$[x_1 + A, x_2 + A, \ldots, x_n + A; y_1 + A, y_2 + A, \ldots, y_n + A; \omega] = A.$$

This, however, is equivalent to

$$[x, y] \in A,$$
$$[x_1, x_2, \ldots, x_n; y, y_2, \ldots, y_n; \omega] \in A,$$

i.e., to (9). //

In particular, *the Ω-factor-group G/G' of the derived group of an Ω-group G is abelian.*

6. A normal series

$$G = A_0 \supset A_1 \supset A_2 \supset \ldots \supset A_k = O \tag{10}$$

of an Ω-group G (see § 20.4) is called a *central series* if

$$[A_i, G] \subseteq A_{i+1} \qquad (i = 0, 1, \ldots, k-1). \tag{11}$$

Note that (11) implies the inclusion

$$[A_i, G] \subseteq A_i,$$

for all i, so that, by § 21.3, all the A_i are ideals in G; in other words, every central series is an invariant series (see § 20.7).

An Ω-group G is called *nilpotent* if it has at least one central series. Every abelian Ω-group G is nilpotent, because the series $G \supset O$ is a central series in G.

The *lower central chain* of an arbitrary Ω-group G is defined as the descending chain of ideals

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \ldots \supseteq G_i \supseteq \ldots, \tag{12}$$

where

$$G_{i+1} = [G_i, G] \qquad (i = 0, 1, 2, \ldots). \tag{13}$$

Note that G_{i+1} is contained in G_i, provided G_i can be proved to be an ideal.

An Ω-group G is nilpotent if and only if its lower central series (12) *reaches the zero subgroup after a finite number of steps, i.e., $G_k = O$ for some k.*

For if $G_k = O$, then (12) becomes a finite central series, in virtue of (13). Conversely, suppose that an Ω-group G has the central series (10). Then $G_0 = G = A_0$. If it is already proved that $G_i \subseteq A_i$, then by (13) and (11),

$$G_{i+1} = [G_i, G] \subseteq [A_i, G] \subseteq A_{i+1}.$$

Hence it follows that

$$G_k \subseteq A_k = O,$$

that is, $G_k = O$. //

Every Ω-subgroup A of a nilpotent Ω-group G is itself nilpotent.

For suppose that

$$A = A_0 \supseteq A_1 \supseteq A_2 \supseteq \ldots \supseteq A_i \supseteq \ldots$$

is the lower central chain of A. Then

$$A_0 = A \subseteq G = G_0.$$

Suppose it already proved that $A_i \subseteq G_i$. Then

$$A_{i+1} = [A_i, A] \subseteq [G_i, G] = G_{i+1}.$$

From $G_k = O$ it now follows that $A_k = O$. //

Every homomorphic image $H = G\varphi$ of a nilpotent Ω-group G is itself nilpotent.

For let (12) and

$$H = H_0 \supseteq H_1 \supseteq H_2 \supseteq \ldots \supseteq H_i \supseteq \ldots$$

be the lower central chains of the Ω-groups G and H, respectively. Then

$$H_0 = H = G\varphi = G_0\varphi.$$

Suppose it already proved that

$$H_i \subseteq G_i\varphi. \tag{14}$$

If $a' \in H_i$, $b' \in H$, then, by (14), there exist elements $a \in G_i$ and $b \in G$ such that $a\varphi = a'$, $b\varphi = b'$ and hence, by the definition of a homomorphism,

$$[a, b]\varphi = [a', b']. \tag{15}$$

Similarly, for an arbitrary n-ary operation $\omega \in \Omega$ and arbitrary $a_1', a_2', \ldots, a_n' \in H_i$, $b_1', b_2', \ldots, b_n' \in H$ there exist elements $a_1, a_2, \ldots, a_n \in G_i$, and $b_1, b_2, \ldots, b_n \in G$ such that $a_i\varphi = a_i'$, $b_i'\varphi = b_i'$ $(i = 1, 2, \ldots, n)$ and therefore

$$[a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \omega]\varphi = \\ = [a_1', a_2', \ldots, a_n'; b_1', b_2', \ldots, b_n'; \omega]. \tag{16}$$

Since, by a remark made at the end of § 18.13, the image $G_{i+1}\varphi$ of the ideal G_{i+1} is an ideal in H, we have, in virtue of (15) and (16),

$$H_{i+1} = [H_i, H] \subseteq G_{i+1}\varphi.$$

Hence $G_k = O$ implies that $H_k = O$. //

7. A normal series

$$G = A_0 \supset A_1 \supset A_2 \supset \ldots \supset A_k = O \tag{17}$$

of an Ω-group G is called a *solvable series* if all the factors

$$A_i/A_{i+1} \qquad (i = 0, 1, \ldots, k-1)$$

of the series are abelian Ω-groups, that is, by § 21.5, if

$$[A_i, A_i] \subseteq A_{i+1} \qquad (i = 0, 1, \ldots, k-1). \tag{18}$$

An Ω-group G is called *solvable* if it has at least one solvable series.

Every nilpotent Ω-group G is solvable.

For if (10) is a central series of G, then by (11),

$$[A_i, A_i] \subseteq [A_i, G] \subseteq A_{i+1} \qquad (i = 0, 1, \ldots, k-1),$$

are required. //

The *derived chain* of an Ω-group G is the descending chain of Ω-subgroups

$$G = G^{(0)} \supseteq G' \supseteq G'' \supseteq \ldots \supseteq G^{(i)} \supseteq \ldots, \tag{19}$$

where

$$G^{(i+1)} = [G^{(i)}, G^{(i)}] \qquad (i = 0, 1, 2, \ldots). \tag{20}$$

An Ω-group G is solvable if and only if its derived chain (19) *reaches zero after a finite number of steps, i.e., if $G^{(k)} = O$ for some k.*

For if $G^{(k)} = O$, then the chain (19) becomes a finite solvable series in virtue of (20). Conversely, let the Ω-group G have a solvable series (17). Then $G^{(0)} = G = A_0$. If it is already proved that $G^{(i)} \subseteq A_i$, then by (20) and (18)

$$G^{(i+1)} = [G^{(i)}, G^{(i)}] \subseteq [A_i, A_i] \subseteq A_{i+1}.$$

Hence it follows that

$$G^{(k)} \subseteq A_k = O,$$

i.e., that $G^{(k)} = O$. //

Every Ω-subgroup A of a solvable Ω-group G is itself solvable.

For let

$$A = A^{(0)} \supseteq A' \supseteq A'' \supseteq \ldots \supseteq A^{(i)} \supseteq \ldots$$

be the derived chain of A. Then

$$A^{(0)} = A \subseteq G = G^{(0)}.$$

If it is already proved that $A^{(i)} \subseteq G^{(i)}$, then

$$A^{(i+1)} = [A^{(i)}, A^{(i)}] \subseteq [G^{(i)}, G^{(i)}] = G^{(i+1)}.$$

From $G^{(k)} = O$ it now follows that $A^{(k)} = O$. //

Every homomorphic image of a solvable Ω-group G is itself solvable.

It is sufficient to prove this statement for an Ω-factor-group G/A. Since every refinement of a solvable series is itself solvable, by Schreier's Theorem of § 20.5, the normal series $G \supset A \supset O$ in our case has a solvable refinement which, by §§ 18.12 and 18.13, determines a solvable series for G/A. //

8. In groups as well as in non-associative rings, and in particular, in Lie rings, solvability is a substantially weaker property than nilpotence. In the case of associative rings, the situation is different:

Every solvable associative ring is nilpotent.

We observe, first of all, that *if R is an associative ring and n a natural number, then the ideal generated in R by the set of all possible products of n elements of R coincides with the subgroup generated by this set in the additive group of R.* For, this subgroup consists of the zero element and all the sums of a finite number of products of n elements. When we multiply such a sum on the left or on the right by an arbitrary element of R, we obtain a sum of products of $n+1$ elements each. However, when one pair of adjacent factors is replaced in each of the terms of this sum by their product, we again have a sum of products of n elements.

From this it is easy to derive the following two statements.

In an associative ring R, the i-th term $R^{(i)}$ of the derived series coincides with the ideal generated in R by the set of all possible products of 2^i elements of R.

For $i = 1$, we know this from § 21.3. Suppose that it has already been proved for a particular i that $R^{(i)}$ consists of sums of products of 2^i elements each. Then, by (20), $R^{(i+1)}$ is the subgroup of the additive pairs of elements of $R^{(i)}$, i.e., in virtue of the distributive laws, by all possible products of 2^{i+1} elements.

In an associative ring R, the i-th term R_i of the lower central series coincides with the ideal generated in R by the set of all possible products of $i+1$ elements of R.

For $i=1$, this is true. Suppose our statement already proved for a particular i, so that R_i consists of sums of products of $i+1$ elements. Then, by (13) and § 21.1, R_{i+1} is the ideal generated in R by the products of elements of R_i with arbitrary elements of R, on the left or on the right—that is, in virtue of the distributive laws, by all possible products of $i+2$ elements.

Thus, for an associative ring R

$$R^{(i)} = R_{2^i-1} \qquad (i=0, 1, 2, \ldots),$$

so that from $R^{(k)}=O$ it follows that $R_{2^k-1}=O$, as required. //

§ 22. Primitive Classes of Universal Algebras

1. We now return to the study of arbitrary universal algebras. The concept of algebras of equal type (see § 17.5) is far too general to distinguish reasonable classes of algebraic formations. For example, by no means every universal algebra with one binary operation (multiplication), one unary operation (the taking of the inverse element), and one nullary operation (the taking of the unit element) is a group, although it is of the same type as groups. Groups can be distinguished among all these algebras by the fact that the following *identical relations* hold in them: for arbitrary x, y, z

$$(xy)z = x(yz), \quad x\cdot 1 = x, \quad xx^{-1}=1.$$

We wish to define the concept of an identical relation for universal algebras with an arbitrary system Ω of operations.

2. Let X be a non-empty auxiliary set whose elements are called *free elements* and are denoted by x with indices, and also by y, z, etc. Furthermore, with every nullary operation of Ω, if they exist, we associate certain symbols, the *symbols of nullary operations.*

Now we define the concept of a *word*. First of all, all free elements and all symbols of nullary operations are counted as words. If we already know that the expressions $w_1, w_2, \ldots, w_n$ are words, then for an arbitrary n-ary operation $\omega\in\Omega$, where $n\geqq 1$, the formal expression

$$w_1 w_2 \ldots w_n \omega \tag{1}$$

is also taken to be a word.

Thus, every word is a finite expression in which free elements, symbols of nullary operations, and symbols ω for n-ary operations of Ω ($n \geqq 1$) occur with an arbitrary number of repetitions; here the symbol ω of an n-ary operation always 'acts' on the n words preceding it.

Suppose, for example, that Ω contains the ternary operation ω, the unary operations ω' and ω'', and the nullary operation with the symbol 0. Then the expressions

$$[(x\omega')0y\omega]x(y\omega'')\omega \quad \text{and} \quad (xyz\omega)(x\omega')(y\omega')\omega$$

are words. Of course, these words could also be written without any parentheses or brackets:

$$x\omega'0y\omega xy\omega''\omega, \qquad xyz\omega x\omega' y\omega'\omega.$$

We shall call $w_1, w_2, \ldots, w_n$ *subwords* of the word (1). Furthermore, we shall assume that the property of being a subword is transitive, so that subwords of $w_1\ w_2, \ldots, w_n$ are also subwords of (1), etc. In particular, all the free elements and symbols of nullary operations occurring in the expression for (1) are subwords of (1). We shall also count every word as a subword of itself.

It is clear that the concept of a word depends essentially on the system of operations Ω. It stands to reason that in examining universal algebras with a system of operations Ω we shall only use words that have meaning in Ω.

3. Let w_1 and w_2 be words relative to a system of operations Ω. Let $x_1, x_2, \ldots, x_k$ denote all the distinct free elements, and $0_1, 0_2, \ldots, 0_l$ all the symbols of nullary operations that occur in at least one of these words. We shall say that in a universal algebra G with the system of operations Ω, the *identical relation*

$$w_1 = w_2 \tag{2}$$

is satisfied if the equation (2) holds in G when the free elements x_i ($i = 1, 2, \ldots, k$) are replaced by arbitrary, not necessarily distinct, elements $a_i \in G$. The symbols 0_j ($j = 1, 2, \ldots, l$) are of course replaced by those elements of G that are taken in accordance with the corresponding nullary operation in G, and quite generally the operations of Ω are carried out according to the rules of procedure in G.

If A is a set of identical relations, then all universal algebras with the system of operations Ω in which all the identical relations of A are satisfied form a *primitive class* of algebras. We shall denote this primitive class by the same letter A as the set of identical relations defining it.

Thus, groups regarded as algebras with one binary, one unary, and one nullary operation form a primitive class. Abelian groups form a narrower primitive class: here the additional identical relation $xy = yx$ is required to hold. We obtain an even narrower primitive class by requiring the additional identical relation $x^2 = 1$ to hold.

Rings, associative rings, associative-commutative rings, Lie rings, and Jordan rings are also primitive classes of universal algebras.

The whole class of universal algebras of equal type with a common system of operations Ω can also be regarded as a primitive class, viz., for an empty set of identical relations. The class of Ω-groups with a given system of operations is also a primitive class: it is defined by the identical relations that occur in the definition of a group and the relations (1) of § 18.1. At the other extreme, the set of identical relations A may be such that the identical relation $x = y$ is one of its consequences, so that the corresponding primitive class consists of the unique algebra containing only a single element.

4. *Every primitive class of universal algebras contains all the subalgebras and all the homomorphic images of each of its algebras.*

Clearly, if the identical relation (2) holds in an algebra G, then it holds, in particular, for the elements of an arbitrary subalgebra. Suppose, on the other hand, that the identical relation (2) holds in G and that φ is a homomorphism of G onto an algebra H of equal type. We let w_1'', w_2'' denote the elements of H that are obtained from w_1 and w_2 by substituting certain elements $b_i \in H$ $(i = 1, 2, \ldots, k)$ for the free elements x_i $(i = 1, 2, 3, \ldots, k)$. We choose elements a_i $(i = 1, 2, \ldots, k)$ in G for which $a_i \varphi = b_i$ for all i, and we let w_1', w_2' denote the elements of G that are obtained by substituting a_i in place of x_i $(i = 1, 2, \ldots, k)$ in the words w_1 and w_2. It follows from the definition of a homomorphism that

$$w_j' \varphi = w_j'' \qquad (j = 1, 2),$$

and since $w_1' = w_2'$, we also have $w_1'' = w_2''$, as required. //

5. The concept of a word can also be used for other purposes. Let us consider a universal algebra G with a system Ω of operations and let us choose an arbitrary word

$$w = w(x_1, x_2, \ldots, x_n)$$

in these operations, relative to free elements $x_1, x_2, \ldots, x_n$ $(n \geqq 1)$. Let us replace k of these elements $(0 \leqq k \leqq n)$—the elements $x_{n-k+1}, x_{n-k+2}, \ldots, x_n$, say—by certain fixed elements $b_1, b_2, \ldots, b_k \in G$. The expression

$$w(x_1, x_2, \ldots, x_{n-k}, b_1, b_2, \ldots, b_k), \tag{3}$$

so obtained defines an $(n-k)$-ary operation in G, in the following way: The operation assigns to the system of elements $a_1, a_2, \ldots, a_{n-k} \in G$ the uniquely determined element

$$w(a_1, a_2, \ldots, a_{n-k}, b_1, b_2, \ldots, b_k).$$

All the operations that can be obtained in G in this manner are called *derived operations* of G. Of course, distinct expressions of the form (3) may happen to define one and the same derived operation in G. A derived operation is called *principal* if $k=0$, i.e., if no preliminary replacement of some of the free elements by elements of G occurs in the word w. The principal derived operations of an algebra Ω include, in particular, the operations of the system Ω itself.

6. Let two universal algebras be given on a set G, with the systems of operations Ω and Ω', respectively. We shall say that these two algebras define one and the same *algebraic system* on G if every operation of Ω' is derived from an operation of Ω, and vice versa; the system of operations Ω and Ω' themselves are called *equivalent* in G. If G and H are algebras of equal type relative to a system of operation Ω and if we go over in G to an equivalent system of operations Ω', then such a transition does not, in general, have meaning in H: if a derived operation was obtained in G by replacing some of the free elements in a certain word by fixed elements of G, then this operation need not be uniquely determined in H. It follows from this that the concept of a homomorphism refers only to a given system of operations Ω and may become meaningless on going over to an equivalent system of operations Ω'. The same is true for the concept of a subalgebra: we only have to bear in mind that the elements $b_1, b_2, \ldots, b_k$ that occur in the expression (3) need not belong to the subalgebra in question (relative to the operations of Ω).

The reader will have no difficulty in verifying that the situation is different if every operation of each of the systems Ω, Ω' is a *principal* derived operation of the other system. Hence it follows that *the two methods of regarding groups as universal algebras that were indicated in* § 17.3 *are, in fact, of equal standing from the point of view of homomorphisms and subgroups,* a fact with which we are already familiar.

§ 23. Free Universal Algebras

1. The set of all possible words relative to a system of operations Ω and a set of free elements X can be regarded as an algebra with Ω as system of operations. For if an n-ary operation $\omega \in \Omega$ $(n \geqq 1)$ and words $w_1, w_2,$

$w_3, \ldots, w_n$ are given, then the word $w_1 w_2 \ldots w_n \omega$ is taken to be the result of applying the operation ω to the given words. The application of any nullary operation of Ω is to be understood as the choice of the symbol of this nullary operation which, as we know, belongs to the set of words.

We denote this *algebra of words* by $S(\Omega, X)$. Of course, no non-trivial identical relation of the form

$$w_1 = w_2, \tag{1}$$

where w_1 and w_2 are distinct words, holds in this algebra. It is also clear that X is a system of generators for this algebra (see § 17.4). If X' is a subset of X, then the subalgebra generated by X' in $S(\Omega, X)$ is the algebra of words $S(\Omega, X')$.

2. Let us now discuss an arbitrary system of identical relations A. Two words v' and v'' are called *equivalent* relative to A if we can pass from one to the other by a finite number of transformations of the following type: suppose that the relation (1) is contained in A; we replace the free elements x_i $(i = 1, 2, \ldots, k)$ occurring in it by certain words, and the left-hand and right-hand sides of (1) now become $\overline{w}_1$ and $\overline{w}_2$; if $\overline{w}_1$ (or $\overline{w}_2$) is a subword of v', then it is replaced in v' by $\overline{w}_2$ (or $\overline{w}_1$, respectively).

This relation is obviously reflexive, transitive, and symmetric, so that we have introduced an equivalence relation in $S(\Omega, X)$. *This is even a congruence* (see § 17.7), because if we have a word $w_1 w_2 \ldots w_n \omega$ obtained from $w_1, w_2, \ldots, w_n$ by applying the operation ω, then the replacement of the words w_i $(i = 1, 2, \ldots, n)$ by words equivalent to them can be effected by applying a finite number of transformations of the type described above to the given word.

The factor algebra of $S(\Omega, X)$ with respect to the congruence so constructed (see § 17.7), which we denote by the symbol $S(\Omega, X, A)$, is called the *free algebra* of the primitive class A, and the set X is called a *system of free generators* of this algebra. Of course, the system of generators is not, in fact, the set X, but the set of corresponding classes of equivalent words; we denote it, nonetheless, by the same letter X. Every algebra isomorphic to $S(\Omega, X, A)$ is also called free, and the set of images of the elements of X in this isomorphism is called a system of free generators.

It follows from the definition of a free algebra $S(\Omega, X, A)$ that all the identical relations of A hold in it, so that *it belongs to the primitive class* A.

3. *Let M be a system of generators in the universal algebra G of the primitive class A. G is a free algebra in the primitive class A, and M a system of free generators of G if and only if the following conditions are satisfied: For every algebra H of the class A and every single-valued mapping φ of M*

into H there exists a unique homomorphic mapping of G into H that coincides with φ on M.

Let us assume first of all that $G = S(\Omega, X, \Lambda)$, that $M = X$, and that a mapping φ of X into H is given. We define a mapping $\bar{\varphi}$ of the whole algebra $S(\Omega, X, \Lambda)$ into H in the following way: $\bar{\varphi}$ coincides with φ on X; $\bar{\varphi}$ maps the symbols of nullary operations into those elements of H the choice of which is indicated by the corresponding nullary operation in H; finally, if the images of the words $w_1, w_2, \ldots, w_n$ under $\bar{\varphi}$ are already uniquely determined and the operation $\omega \in \Omega$ is n-ary, then we set

$$(w_1 w_2 \ldots w_n \omega)\bar{\varphi} = (w_1\bar{\varphi})(w_2\bar{\varphi}) \ldots (w_n\bar{\varphi})\omega; \qquad (2)$$

it is clear that if we wish this mapping to be a homomorphism, then the image of the word $w_1 w_2 \ldots w_n \omega$ under the mapping $\bar{\varphi}$ can be defined only by (2).

Note that the application of $\bar{\varphi}$ to a word equivalent to $w_1 w_2 \ldots w_n \omega$ in the sense of § 23.2 gives an element that is equal in H to the right-hand side of (2), because all the identical relations of Λ hold in H. Thus, the mapping $\bar{\varphi}$ is defined for all the elements of $S(\Omega, X, \Lambda)$; that it is a homomorphism and uniquely determined follows from (2).

Now let G be an algebra and M a system of its generators satisfying the condition of our theorem. For H we choose a free algebra $S(\Omega, X, \Lambda)$ such that the sets M and X are equipotent, that is, such that there exists a one-to-one mapping φ of M onto X. By hypothesis, there exists a homomorphism $\psi\colon G \to S(\Omega, X, \Lambda)$ that coincides with φ on M. On the other hand, from what has been shown above it follows that there exists a homomorphism $\chi\colon S(\Omega, X, \Lambda) \to G$ that coincides on X with φ^{-1}.

The product $\chi\psi$ is an endomorphism of the algebra $S(\Omega, X, \Lambda)$ that is the identity on the system of generators X, and the product $\psi\chi$ is an endomorphism of the algebra G that is the identity on the system of generators M. By the uniqueness condition, which is assumed for G and has already been proved for $S(\Omega, X, \Lambda)$, it follows that these endomorphisms are the identity automorphisms of the respective algebras. Hence it follows that each of the homomorphisms ψ and χ is, in fact, an isomorphism between the algebras G and $S(\Omega, X, \Lambda)$, as required.

4. A consequence of this theorem is the following result, which illustrates the true significance of the concept of a free algebra:

Every algebra G of the primitive class Λ is a homomorphic image of a free algebra of the same class.

For let us choose in G an arbitrary system of generators M and let us examine a free algebra $S(\Omega, X, \Lambda)$ for which there exists a mapping φ of

X onto M, so that the cardinal number of X is greater than or equal to that of M. Then, by what has been proved, the mapping φ can be extended to a homomorphism $\bar{\varphi}: S(\Omega, X, \Lambda) \to G$; and this is a homomorphism onto the whole of G, because the image of the algebra $S(\Omega, X, \Lambda)$ under $\bar{\varphi}$ contains the whole set M. //

5. It is clear that if the sets X and Y are equipotent, then the free algebras $S(\Omega, X, \Lambda)$ and $S(\Omega, Y, \Lambda)$ are isomorphic. As to the question whether, conversely, an isomorphism of these free algebras allows us to deduce that the set of free generators X and Y are equipotent, the answer, in general, is in the negative. This is shown by the primitive class consisting of a single one-element algebra, which we mentioned in § 22.3; this class in which, of course, arbitrary identical relations hold, we shall call *absolutely degenerate*.

The following theorems are due to Fujiwara [*Proc. Japan. Acad.*, vol. 31 (1955), pp. 135-36].

If a primitive class Λ is not absolutely degenerate and if the free algebras $S(\Omega, X, \Lambda)$ and $S(\Omega, Y, \Lambda)$ are isomorphic and at least one of the sets X, Y is infinite, then these sets are equipotent.

Suppose that this is not so and that the cardinal number of, say, X is greater than that of Y.[1] We may assume that the algebras in question not only are isomorphic but actually coincide:

$$S(\Omega, X, \Lambda) = S(\Omega, Y, \Lambda) = S. \tag{3}$$

Every element of Y can therefore be written as a word in a finite number of free generators from X. When we fix one such expression for each $y \in Y$ and collect all the elements of X that are used in these expressions for all the $y \in Y$, we obtain a subset $X' \subset X$ that necessarily has a smaller cardinal number than X: it is finite if the set Y is finite and has a cardinal number not exceeding that of Y if Y is infinite. We may therefore choose an element $x_0 \in X \setminus X'$.

The element x_0 can be written as a word in a finite number of elements from Y. When we now express these elements, in turn, by means of elements from X', we find that x_0 is equal in S to a certain word w_0 in elements from X',

$$x_0 = w_0.$$

This equation signifies the equivalence of x_0 and w_0 in the algebra of words (see § 23.2), i.e., it can be realized as an application of the identical relations

[1] Here we make use of the result, which is based on the Axiom of Choice, that any two cardinal numbers can be compared in magnitude.

from A. But then, when we bear in mind that the free element x_0 does not occur in the word w_0, we could use the same identical relations to obtain, in S, the equation $a = w_0$, where a is an arbitrary element of S. Thus, in S all elements are equal to each other, that is, the identical relations A imply the identical relation $x = y$, so that the primitive class A turns out to be absolutely degenerate, contrary to assumption. //

6. *Suppose that the system of identical relations A can be extended to a system A^* such that the primitive class A^* is not absolutely degenerate, but that finite sets of free generators generate finite free algebras in it. Then isomorphic free algebras in A have equipotent systems of free generators.*

For let X and Y be systems of free generators of two isomorphic free algebras in A. As above, we may assume that these algebras coincide, i.e., that (3) holds. Furthermore, in virtue of the preceding theorem, both the sets X and Y may be taken to be finite. Let us assume that X contains more elements than Y. We impose an identical relation from the system $A^* \setminus A$ on the algebra S, that is, we go over to the factor algebra S^* with respect to the corresponding congruence. This is obviously a free algebra in the primitive class A^*; and by the conditions of the theorem it is finite, because it has the finite sets X and Y as systems of free generators.

Let the subset $X' \subset X$ contain as many elements as Y. Then the subalgebra $\{X'\}$ of S^*, being a free algebra $S(\Omega, X', A^*)$, must in fact consist of as many elements as S^*, which means that it coincides with the whole algebra S^*. Hence it follows that every element $x_0 \in X \setminus X'$ can be expressed as a word in elements from X'; however, as in the proof of the preceding theorem, this leads to a contradiction. //

7. The class of *groupoids* is a primitive class relative to a single binary operation and an empty set of identical relations; as usual, we shall call this operation multiplication and shall write it in the form ab. The *free groupoid* with a set X of free generators coincides with the groupoid of words. A *word* in this case is a finite ordered system of elements from X,

$$x_1 x_2 \dots x_n \qquad (n \geqq 1),$$

with arbitrary repetitions and with a given *distribution of parentheses*: each of the symbols x_i $(i = 1, 2, \dots, n)$ is to be thought of as enclosed in parentheses, and then further parenthees are inserted in such a way that only two parentheses are 'multiplied' each time. The *product* of two words means that the given words are put into parentheses and then written in juxtaposition.

Of course, the words in groupoids are written as

$$((x_1x_2)x_1)(x_3x_2),$$

for example, and not as

$$((((x_1)(x_2))(x_1))((x_3)(x_2))).$$

8. The class of *semigroups* is a primitive class with the identical relation of associativity. In this case, a *word* relative to a set X of free generators can be written in the form

$$x_1x_2\ldots x_n \qquad (n \geqq 1) \tag{4}$$

without any parentheses. The *product* of two words (4) now means that the given words are written down in juxtaposition, without parentheses.

The statement that we have constructed a *free semigroup* with the set X of free generators will only be justified after we have shown that distinct words of the form (4) represent distinct elements in this semigroup. However, the multiplication of words defined in the preceding paragraph obviously converts the set of all words of the form (4) into a semigroup with X as a set of generators. This semigroup is a homomorphic image of the free semigroup with X as a set of free generators, and since distinct words of the form (4) are distinct elements in it, this is also true for the free semigroup. Hence the homomorphism in question is certainly an isomorphism.

The use of powers of elements enables us to write words in semigroups concisely: instead of the word $x_1x_1x_1x_2x_1x_2x_2$, we can write $x_1^3x_2x_1x_2^2$.

9. Let us now examine the primitive class of all *groups* with the operations of multiplication, of taking inverses, and of taking the unit element. The identical relations enable us to write every *word* in a set X of free generators either in the form 1—this word, containing no free elements at all, is called the *empty* word—or else in the form

$$w = x_1^{\varepsilon_1}x_2^{\varepsilon_2}\ldots x_n^{\varepsilon_n} \qquad (n \geqq 1), \tag{5}$$

where the x_i $(i = 1, 2, \ldots, n)$ are elements of X, not necessarily distinct, $\varepsilon_i = \pm 1$ $(i = 1, 2, \ldots, n)$, and if $x_i = x_{i+1}$ then $\varepsilon_i = \varepsilon_{i+1}$; in other words, an element $x \in X$ and its inverse x^{-1} cannot occur side by side in the expression (5). We call the number n the *length* of the word (5).

Distinct words of the form (5) *are distinct elements of the free group with X as set of free generators and are not equal to the empty word in this group.*

To prove this, we construct a group whose elements are the empty word and all possible words of the form (5). We define multiplication

of words in accordance with the way this was done in the free group. The empty word shall play the role of the unit element. If two words of the form (5) are given,

$$w_1 = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_n^{\varepsilon_n}, \qquad w_2 = y_1^{\eta_1} y_2^{\eta_2} \dots y_m^{\eta_m},$$

where $y_i \in X$ $(i = 1, 2, \dots, m)$, then we write these words next to one another:

$$x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_n^{\varepsilon_n} y_1^{\eta_1} y_2^{\eta_2} \dots y_m^{\eta_m}.$$

Now if $x_n = y_1$ and $\varepsilon_n = -\eta_1$ in this expression, then we carry out a *cancellation*. Afterwards we have the elements x_{n-1} raised to the exponent ε_{n-1} and y_2 raised to the exponent η_2 next to each other, and possibly we can carry out yet another cancellation. We continue in this way until we come to the first place k for which either $x_{n-k} \neq y_{k+1}$, or $x_{n-k} = y_{k+1}$ and $\varepsilon_{n-k} = \eta_{k+1}$. Then we obtain

$$w_1 w_2 = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_{n-k}^{\varepsilon_{n-k}} y_{k+1}^{\eta_{k+1}} y_{k+2}^{\eta_{k+2}} \dots y_m^{\eta_m}. \tag{6}$$

Clearly, the right-hand side of (6) is a word of the form (5) if $n \neq m$, or if $n = m$ and $k < n$; otherwise, it is the empty word. Therefore we find that the *inverse* of (5) is the word

$$w^{-1} = x_n^{-\varepsilon_n} \dots x_2^{-\varepsilon_2} x_1^{-\varepsilon_1}. \text{ //}$$

The multiplication of words introduced above is associative.

We shall prove the equation

$$w_1(w_2 w_3) = (w_1 w_2) w_3 \tag{7}$$

by induction on the length of the word w_2, and we may assume that none of the words w_1, w_2, w_3 is empty.

First let $w_2 = y^\eta$, where $y \in X$, $\eta = \pm 1$. If, of the last symbol of w_1 and the first symbol of w_3, at least one is different from $y^{-\eta}$, then there are no cancellations in at least one of the products w, w_2, $w_2 w_3$, and therefore (7) holds. But if

$$w_1 = x_1^{\varepsilon_1} \dots x_{n-1}^{\varepsilon_{n-1}} y^{-\eta} \qquad w_3 = y^{-\eta} z_2^{\delta_2} \dots z_s^{\delta_s},$$

then

$$x_1^{\varepsilon_1} \dots x_{n-1}^{\varepsilon_{n-1}} y^{-\eta} z_2^{\delta_2} \dots z_s^{\delta_s},$$

is a word of the form (5) that is equal both to the left-hand and right-hand side of (7).

Now let $w_2 = w_2' \cdot y_m^{\eta_m}$, where the word w_2' has the form

$$w_2' = y_1^{\eta_1} \dots y_{m-1}^{\eta_{m-1}}.$$

Assuming that the equation (7) is already proved for the case in which the length of w_2 is less than m, we obtain

$$\begin{aligned} w_1(w_2w_3) &= w_1[(w_2' \cdot y_m^{\eta_m})w_3] = w_1[w_2'(y_m^{\eta_m}w_3)] = \\ &= (w_1w_2')(y_m^{\eta_m}w_3) = [(w_1w_2')y_m^{\eta_m}]w_3 = \\ &= [w_1(w_2'y_m^{\eta_m})]w_3 = (w_1w_2)w_3, \end{aligned}$$

as required. //

Thus, we have constructed a group from the words of the form (5) and the empty word for which the set X is a system of generators. This group is a homomorphic image of the free group with X as set of free generators. Hence the words in question are distinct in the free group, as required. *Therefore the group constructed above proves to be the free group with X as system of free generators.*

The words of the form (5) can, of course, be written concisely, as in the case of semigroups, by using powers of elements, but this time negative powers as well.

The free group with a single free generator x is, obviously, the infinite cyclic group $\{x\}$. If the set of free generators X contains more than one element, then the free group generated by X is non-commutative, because the words xy and yx $(x, y \in X)$ are distinct elements in this free group.

In § 24.12 we shall give an account of the subgroups of a free group.

10. In the primitive class of all *abelian groups* written additively, *words* in a set X of free generators are sums of a finite number of distinct elements from X, taken with certain non-zero integer coefficients, and also, the empty word 0. *Addition* of such words is defined as addition of the coefficients of equal elements $x \in X$; but if the element x is absent from one of the words, then we take its coefficient to be zero. A direct verification shows that we obtain an abelian group, which is the free abelian group with X as set of free generators.

11. Let us now proceed to a discussion of the primitive class of all *rings*. Arguing as in the case of semigroups or groups, we find that the additive group of a *free ring* with X as a set of free generators is a free abelian group (see § 23.10) in a set of free generators $\overline{X}$, which in turn is the free multiplicative groupoid with X as a set of free generators (see § 23.7). Multiplication in a free ring reduces, by virtue of the distributive laws, to

the multiplication of elements from $\overline{X}$, which in turn reduces to the law of multiplication of words in a free groupoid (see § 23.7).

Everything we have said remains valid for the *free associative ring* with X as a set of free generators. However, in this case the set $\overline{X}$ is the free semigroup with X as set of free generators (see § 23.8), and multiplication of elements of $\overline{X}$ is carried out according to the law of multiplication of words in a free semigroup.

Finally, the *free associative-commutative ring* with X as a set of free generators is simply the ring of polynomials in elements of X with integer coefficients.

We remark that by no means every primitive class of universal algebras admits a unique ('canonical') expression for the elements of free algebras of this class such as we have obtained in the cases discussed above.

12. *Free abelian groups (and also free groups, free rings, free associative, and free associative-commutative rings) are isomorphic if and only if their systems of free generators are equipotent.*

For if we impose the identical relation $2x = 0$ on an abelian group, we obtain the primitive class of those abelian groups whose non-zero elements are all of order 2. We choose a finite system of free generators X consisting of the elements $x_1, x_2, \ldots, x_n$. Then a word, in the case in question, is an expression of the form

$$k_1x_1 + k_2x_2 + \ldots + k_nx_n,$$

where all the k_i $(i = 1, 2, \ldots, n)$ are equal to one or zero. Addition of words is carried out as addition of coefficients of equal elements x_i reduced modulo 2, so that $1 + 1 = 0$. Thus, we see that in this primitive class the free groups generated by finite systems of free generators are themselves finite.

We arrive at the same primitive class by imposing on the class of groups the identical relations $xy = yx$ and $x^2 = 1$ (actually, it is easy to see that the first of these is a consequence of the second). Essentially the same primitive class is obtained by imposing on the rings mentioned in the statement of the theorem the identical relations

$$xy = 0 \qquad \text{and} \qquad 2x = 0.$$

To prove the theorem, it now remains to refer to the second theorem of Fujiwara (§ 23.6). //

13. Let us consider an arbitrary primitive class of Ω-groups. Every Ω-group G of this class is, by § 23.4, a homomorphic image of a certain

free Ω-group S of the same class, for by §§ 18.6 and 17.8 it is isomorphic to an Ω-factor-group of the free Ω-group S with respect to a certain ideal A. Let the elements s_i, where i ranges over an index set I, generate the ideal A in S. The corresponding elements in G are zero. The system of equations

$$s_i = 0 \qquad (i \in I), \tag{8}$$

where the s_i are assumed to be expressed in terms of free generators of the free Ω-group S, completely determines the ideal A in the free Ω-group S; that is, it determines the Ω-factor-group S/A, and hence it determines the original Ω-group G to within isomorphism. The system of equations (8) is called *a system of defining relations* for the Ω-group G.

Thus, every Ω-group—in particular, every group and every ring—can be given by a system of defining relations in a certain system of generators. This representation is of course by no means unique.

Example: The symmetric group S_3 of degree 3 (see § 7.8) can be given in the class of groups by two generators a, b and the defining relations

$$a^3 = 1, \quad b^2 = 1, \quad abab = 1. \tag{9}$$

For, the permutations

$$a = \begin{pmatrix} 123 \\ 231 \end{pmatrix}, \qquad b = \begin{pmatrix} 123 \\ 213 \end{pmatrix}$$

generate the whole group S_3 and satisfy the relations (9). On the other hand, it follows from (9) that $ba = a^2b$, so that every element of the group given by the defining relations (9) can be written in the form $a^k b^l$, where $k = 0, 1, 2$ and $l = 0, 1$, and hence the group consists of not more than six elements; but the order of S_3 is six.

§ 24. Free Products of Groups

1. The characteristic property of free algebras of a given primitive class $\mathfrak{A}$ mentioned in § 23.3 suggests the following definition. If A_i $(i \in I)$ is a collection of algebras of the class $\mathfrak{A}$, then we call an algebra G of the class $\mathfrak{A}$ the *free union* of this collection if $A_i \subset G$ $(i \in I)$ and if for every algebra H of $\mathfrak{A}$ and every choice of homomorphisms φ_i of the A_i into H $(i \in I)$ there exists a unique homomorphism φ of G into H that coincides with φ_i on A_i $(i \in I)$.

If both the algebras G and G' are free unions of the algebras A_i $(i \in I)$, then there exists an isomorphism between G and G' that extends the identity automorphisms of the subalgebras A_i $(i \in I)$.

For these identity automorphisms can be regarded as homomorphisms φ_i of the subalgebras A_i of G $(i \in I)$ into G', and also as homomorphisms ψ_i of the subalgebras A_i of G' $(i \in I)$ into G. Therefore they induce homomorphisms $\varphi: G \to G'$ and $\psi: G' \to G$. The product $\varphi\psi$ is then the unique homomorphism of G into itself that is determined by the isomorphic embeddings of all the A_i in G as subalgebras, that is, it is the identity automorphism of G. Similarly, $\psi\varphi$ is the identity automorphism of the algebra G'; and therefore the isomorphisms φ and ψ are inverse to each other. //

The free union G of the algebras A_i $(i \in I)$ coincides with the subalgebra generated in G by all the A_i $(i \in I)$.

For let G_0 be the subalgebra generated by the A_i $(i \in I)$ in G (see § 17.4). If an algebra H of the class $\mathfrak{A}$ is given and homomorphisms $\varphi_i: A_i \to H$ $(i \in I)$, then the homomorphism $\varphi: G \to H$ determined by these homomorphisms induces a homomorphism $\varphi_0: G_0 \to H$ that coincides with φ_i on the subalgebra A_i $(i \in I)$. This is the unique homomorphism of G_0 into H with the requisite properties. For it is easy to verify that the subalgebra G_0 consists of precisely those elements of G that can be expressed in at least one way in terms of elements of the subalgebras A_i $(i \in I)$ by means of the operations of G applied a finite number of times. Therefore the homomorphisms φ_i $(i \in I)$ can only generate one homomorphism of the algebra G_0.

Thus, G_0 itself turns out to be the free union of the algebras A_i $(i \in I)$. But then, as we have shown, there exists an isomorphism φ of G_0 onto G generated by the identity automorphisms of the algebras A_i $(i \in I)$. By what has been proved in the preceding paragraph, these identity automorphisms can only generate the identity automorphism of G_0, so that $G_0 = G$. //

2. A group G is called the *free product* of its subgroups A_i $(i \in I)$ if these subgroups generate the whole group G and if every element $g \in G$, other than 1, has a *unique* decomposition into a product of the form

$$g = a_1 a_2 \dots a_n \qquad (n \geqq 1), \tag{1}$$

where $a_k \in A_{i_k}$ and $a_k \neq 1$ $(k = 1, 2, \dots, n)$ and where adjacent elements do not belong to one and the same subgroup A_i, so that

$$i_k \neq i_{k+1} \qquad (k = 1, 2, \dots, n-1).$$

The product with these properties on the right-hand side of (1) is called the *irreducible representation* of g. The number n is called the *length* of this element and is denoted by $\lambda(g)$. The indices i_1 and i_n will be called the *first* and the *last index* of g, respectively. We also put $\lambda(1) = 0$.

As a notation for the free product we shall use the symbol

$$G = \prod_{i \in I}^{*} A_i$$

or, if the number of free factors is finite,

$$G = A_1 * A_2 * \ldots * A_n.$$

A comparison with § 23.9 shows that *a free group is the free product of infinite cyclic groups.*

If a group G is the free product of subgroups A_i $(i \in I)$, then G is the free union of the groups A_i in the primitive class of groups.

Indeed, if for a certain choice of homomorphisms φ_i of the subgroups A_i into a group H $(i \in I)$ there exists a homomorphism φ of G into H such that $\varphi = \varphi_i$ on A_i $(i \in I)$, then by (1) we have for all $g \in G$

$$g\varphi = a_1\varphi_{i_1} \cdot a_2\varphi_{i_2} \cdot \ldots \cdot a_n\varphi_{i_n}, \tag{2}$$

so that φ is uniquely determined. It is easy to verify, on the other hand, that if (2) is regarded as a definition of the mapping φ of G into H, then φ is a homomorphism. //

This theorem has a converse; this follows immediately from results of the present and the preceding subsections in virtue of a theorem to be proved in the following subsection.

3. We now show that *for every collection of groups A_i $(i \in I)$, the free product of this collection of groups exists* or, more precisely, the free product of groups isomorphic to the given ones.

We define a *word* as a formally written expression

$$a_1 a_2 \ldots a_n \qquad (n \geqq 1), \tag{3}$$

where $a_k \in A_{i_k}$, $a_k \neq 1$ $(k = 1, 2, \ldots, n)$ and where $i_k \neq i_{k+1}$ $(k = 1, 2, 3, \ldots, n-1)$.

We also introduce the *empty word* (the case $n = 0$). We denote the set of all words by M and the symmetric group on this set (see § 7.8) by S.

Every element $a \in A_i$ determines a mapping $\bar{a}$ of M into itself in the following way: it carries the empty word into the word a of length 1; from this it follows that for $a \neq b$ $(a, b \in A_i)$ we have $\bar{a} \neq \bar{b}$. Furthermore, for words of the form (3), we set

$$(a_1 a_2 \ldots a_n)\bar{a} = \begin{cases} a_1 a_2 \ldots a_n a & \text{if } i_n \neq i; \\ a_1 a_2 \ldots (a_n a) & \text{if } i_n = i \text{ and } a_n a \neq 1; \\ a_1 a_2 \ldots a_{n-1} & \text{if } i_n = i \text{ and } a_n a = 1. \end{cases}$$

Hence it follows, in particular, that the unit element of A_i corresponds to the identity mapping of M into itself.

It is easy to verify that

$$\bar{a}\bar{b} = \overline{ab} \qquad (a, b \in A_i),$$

where the product on the right-hand side is to be understood in the sense of multiplication of mappings, so that $\overline{a^{-1}}$ is the inverse mapping of $\bar{a}$. Thus, every mapping $\bar{a}$ is a permutation of M and the mappings $\bar{a}$ taken for all $a \in A_i$ therefore form, in S, a subgroup $\bar{A}_i$ isomorphic to A_i.

Let G denote the subgroup of S generated by all the subgroups $\bar{A}_i$ $(i \in I)$. The group G is the free product of the subgroups $\bar{A}_i$ $(i \in I)$: every element of G other than the unit element can be written in the form of a product

$$\bar{a}_1 \bar{a}_2 \dots \bar{a}_n, \tag{4}$$

where $a_1 a_2 \dots a_n$ is a word; and this expression is unique, because the permutation (4) carries the empty word exactly into $a_1 a_2 \dots a_n$. //

4. Let us now prove the following *Subgroup Theorem* for a free product of groups [A. G. Kurosh, *Math. Ann.*, vol. 109 (1934), pp. 647-60]:

If

$$G = \prod_{i \in I}{}^{*} A_i, \tag{5}$$

then every subgroup U of G is a free product of subgroups conjugate in G (see § 13.4) *to certain subgroups of the free factors A_i, and of a certain free subgroup.*

In order to give a more explicit formulation of this theorem, we introduce the following concept. Two elements $x, y \in G$ are called *equivalent with respect to the double module* (A_i, U) if

$$y = a_i x u \qquad (a_i \in A_i;\ u \in U).$$

Since A_i and U are subgroups, we obtain a partition of G into disjoint classes of the form $A_i x U$, which we call a *decomposition* (A_i, U).

Observe that if D is a class of the decomposition (A_i, U), then, for every $x \in D$, D contains the whole left coset xU, so that it consists of several left cosets modulo U.

The formulation of the subgroup theorem in fuller form is as follows:

If (5) *holds and $U \subset G$, then, in every class D of each of the decompositions (A_i, U) $(i \in I)$, exactly one representative*

$$s = s(i, D)$$

can be chosen such that

$$U = \prod_{i \in I,\, D \in (A_i,\, U)}^{*} (U \cap s^{-1} A_i s) * F, \tag{6}$$

where F is a free subgroup. Moreover, the representatives can be chosen so that for all $i \in I$ the unit element is the representative of its class $A_i U$, i.e., all the intersections $U \cap A_i$ $(i \in I)$ other than E occur as free factors in (6).

5. The proof of this theorem, to be given below, is due to S. MacLane [*Mathematika*, vol. 5 (1958), pp. 13-19].

Lemma 1: *Systems of representatives $r_i(C)$ depending on $i \in I$ can be chosen in all left cosets C of U such that the following conditions are satisfied*:

1) $r_i(U) = 1$;

2) *If $a_i \in A_i$, then $r_i(a_i C) = a_i' r_i(C)$, where $a_i' \in A_i$;*

3) *If $r_i(C) = a_i s$, where $a_i \in A_i$ (here a_i may be equal to the unit element) and $s \neq 1$, and if the first index j of s is different from i, then*

$$r_i(sU) = r_j(sU) = s.$$

We shall make the choice of the representatives $r_i(C)$ by induction on the minimal length $\lambda(A_i C)$ of elements of that class $A_i C$ of the decomposition with respect to the double module (A_i, U) which contains C. We shall also assume that the following condition is satisfied:

4) $\lambda(r_i(C)) \leqq 1 + \lambda(A_i C)$.

If $\lambda(A_i C) = 0$, i.e., $A_i C = A_i \cdot 1 \cdot U = A_i U$, then C contains elements of A_i and we choose any one of these elements for $R_i(G)$, provided only that $r_i(U) = 1$. The conditions 1), 2), and 4) are verified without difficulty, and 3) is vacuous in this case.

Suppose now that the class D of the decomposition (A_i, U) is such that $\lambda(D) = n \geqq 1$. If $g \in D$ and $\lambda(g) = n$, then $D = A_i g U$ and the first index j of g is different from i, because otherwise the class D would contain an element whose length is less than n. Since $\lambda(A_j g U) < n$, the representative $r_j(gU) = s$ is already chosen, by the induction hypothesis; and by 4), $\lambda(s) \leqq n$. Since $D = A_i s U$, we have $n = \lambda(D) \leqq \lambda(s)$; and consequently $\lambda(s) = n$, so that the first index k of s is different from i and, by the induction hypothesis (condition 3)), $r_k(sU) = s$.

We now choose the representatives $r_i(C)$ for all C that occur in D. For if $C = sU$, then we put $r_i(C) = s$; for every other $C \subseteq D$ we choose as $r_i(C)$ one of the elements $a_i s$ $(a_i \in A_i)$ contained in C. It is clear that the conditions 2), 3), and 4) hold and 1) is not applicable to the case $n \geqq 1$. //

The lemma is now proved. At the same time we have seen that, in each class $D \in (A_i, U)$, among the representatives $r_i(C)$ of all left cosets C occurring in D there is exactly one representative s that is either equal to 1 or has its first index different from i. For this is the representative chosen as $s(i, D)$. We mention once more that in every class C occurring in D the representative $r_i(C)$ differs from $s(i, D)$ by a left factor belonging to A_i.

6. LEMMA 2: *The subgroup U is generated by all the elements $r_i^{-1}(C)r_j(C)$, where $i, j \in I$, and C ranges over the left cosets of U, and by all the intersections $U \cap (s^{-1}A_i s)$, where $s = s(i, D)$ $(D \in (A_i, U); i \in I)$.*

First of all, we observe that both the representatives $r_i(C)$ and $r_j(C)$ belong to C, that is, they differ from each other by right factors from U, so that

$$r_i^{-1}(C)r_j(C) \in U.$$

We now take an arbitrary element $g \in G$, an arbitrary expression (not necessarily irreducible) for g in the form

$$g = a_1 a_2 \dots a_n \qquad (a_k \in A_{i_k};\ k = 1, 2, \dots, n), \tag{7}$$

and an arbitrary left coset C of U. Then a simple verification, which is most conveniently carried out from the bottom up, shows the validity of the following equation, on the right-hand side of which is the product of the elements in brackets:

$$\begin{aligned} r_{i_1}^{-1}(gC) \cdot g \cdot r_{i_n}(C) &= r_{i_1}^{-1}(gC) \cdot a_1 a_2 \dots a_n \cdot r_{i_n}(C) = \\ &= [r_{i_1}^{-1}(gC) \cdot a_1 \cdot r_{i_1}(a_2 \dots a_n C)] \cdot \dots \cdot [r_{i_{n-2}}^{-1}(a_{n-1}a_n C)] \cdot r_{i_{n-1}}(a_{n-1}a_n C)] \cdot \\ &\quad \cdot [r_{i_{n-1}}^{-1}(a_{n-1}a_n C) \cdot a_{n-1} \cdot r_{i_{n-1}}(a_n C)] \cdot \\ &\quad \cdot [r_{i_{n-1}}^{-1}(a_n C) \cdot r_{i_n}(a_n C)] \cdot [r_{i_n}^{-1}(a_n C) \cdot a_n \cdot r_{i_n}(C)]. \end{aligned}$$

Factors of the form

$$r_i^{-1}(C')r_j(C').$$

alternate with other factors; let us examine any of these other factors

$$x = r_{i_k}^{-1}(a_k a_{k+1} \dots a_n C) \cdot a_k \cdot r_{i_k}(a_{k+1} \dots a_n C) \quad (k = 1, 2, \dots, n)$$

and show that it is contained in the intersection $U \cap s^{-1}A_{i_k}s$, where $s = s(i_k, D)$, $D \in (A_{i_k}, U)$, $C' \subseteq D$, and

$$C' = a_{k+1} \dots a_n C.$$

For since

$$r_{i_k}(a_k C') = a_{i_k} s \quad \text{and} \quad r_{i_k}(C') = a'_{i_k} s,$$

where $a_{i_k}, a'_{i_k} \in A_{i_k}$, we have

$$x = s^{-1}(a_{i_k}^{-1} a_k a'_{i_k}) s \in s^{-1} A_{i_k} s.$$

On the other hand, if $C = cU$, then

$$r_{i_k}(a_k a_{k+1} \dots a_n C) = a_k a_{k+1} \dots a_n c u_1 \quad (u_1 \in U),$$
$$r_{i_k}(a_{k+1} \dots a_n C) = a_{k+1} \dots a_n c u_2 \quad (u_2 \in U).$$

and therefore $x = u_1^{-1} u_2 \in U$.

We now observe that if $C = U$ and $g \in U$, then

$$r_{i_1}(gC) = r_{i_n}(C) = 1,$$

so that the fundamental equation obtained above reduces to the expression for g, thus proving our lemma.

7. Let us now introduce symbols $[C, i, j]$ for arbitrary $i, j \in I$ and arbitrary left cosets of C in U. Let F denote the group that has the set of all these symbols as a system of generators and is given in these generators by the system of defining relations (see § 23.13) consisting of all possible equations of the following types:

(a) $[C, i, j][C, j, k] = [C, i, k]$,

(b) $[U, i, j] = 1$,

(c) $[sU, i, j] = 1$ if $s = s(i, D)$ and j is the first index of s.

Lemma 3: *F is a free group.*

For the equation

$$[C, i, j] = [C, i, k] \cdot [C, j, k]^{-1}$$

follows from (a), so that only the symbols $[C, i, k]$ with a fixed second index k need be retained in the system of generators. The defining relations in these generators are the equations of the following types:

(a′) $[C, k, k] = 1$,

(b′) $[U, i, k] = 1$,

(c′) $[sU, i, k] = [sU, j, k]$, where $s = s(i, D)$ and j is the first index of s.

Finally, if we discard all generators equal to 1 and leave only one representative in each class of equal generators, we obtain a system of generators for F that are not connected by any non-trivial relations, i.e., a system of free generators. //

8. To every subgroup $U \cap s^{-1} A_i s$, where $s = s(i, D), D \in (A_i, U)$ $(i \in I)$ we assign a group $B_{i,D}$ isomorphic to it and we fix a definite isomorphism $\sigma_{i,D}: B_{i,D} \to (U \cap s^{-1} A_i s)$. Furthermore, to every element $[C, i, j]$ of F

we assign the element $r_i^{-1}(C)r_j(C)$ of U. Since after this change the equations (a), (b) and (c) remain inviolate (see 1) and 3) of Lemma 1), we have determined a homomorphism o_F of F into U.

In accordance with § 24.3, we now take the group

$$B = \prod_{i \in I, D \in (A_i, U)}^{*} B_{i,D} * F. \tag{8}$$

By § 24.2, *the set of all homomorphisms $o_{i,D}$ and o_F induces a uniquely determined homomorphism o of B into U*—in fact, by Lemma 2, onto U.

9. We now return to our discussion of the fundamental formula of § 24.6, which gives an expression for the element $r_{i_1}^{-1}(gC) \cdot g \cdot r_{i_n}(C)$. If, on the right-hand side of this formula, every factor of the form $r_i^{-1}(C)r_j(C)$ is replaced by the corresponding element $[C, i, j]$ of F and every factor from a subgroup of the form $U \cap s^{-1}A_i s$ by the element of $B_{i,D}$ corresponding to it under the isomorphism $o_{i,D}$, then we obtain a well-defined element of B, which we denote by $m(g, C)$.

LEMMA 4: *The element $m(g, C)$ does not depend on the choice of the expression (7) for g and has the following properties:*

$$m(gh, C) = m(g, hC) \cdot [hC, i, j] \cdot m(h, C), \tag{9}$$

where i is the last index of g and j is the first index of h;

$$m(1, C) = 1; \tag{10}$$

$$m(g^{-1}, C) = m^{-1}(g, g^{-1}C). \tag{11}$$

All these statements follow easily from the definition of $m(g, C)$. We need only take into account that the element g has a unique irreducible representation in the free decomposition (5) and that we can pass from one expression of the form (7) for g to any other such expression by a finite number of the following transformations and their inverses: If $i_k = i_{k+1}$ and $a_k a_{k+1} = a' \in A_{i_k}$, then the segment $a_k a_{k+1}$ in (7) can be replaced by the element a' when $a' \neq 1$, and can be omitted when $a' = 1$.

The equation (9) for $C = U$ and $g, h \in U$ shows, by (b) of § 24.7, that *the mapping τ that assigns to every element $g \in U$ the element $m(g, U) \in B$ is a homomorphism of U into B.* //

LEMMA 5: *The product τo is the identity mapping of U onto itself.*

For from the definition of the homomorphism o and of the element $m(g, C)$ it follows that

$$m(g, C)o = r_{i_1}^{-1}(gC) \cdot g \cdot r_{i_n}(C), \tag{12}$$

so that for $g \in U$, by 1) of Lemma 1,

$$g(\tau\sigma) = m(g, U)\sigma = g. \text{ //}$$

10. Lemma 6: *If* $r = r_i(C)$, *then* $m(r, U) = 1$.

We shall carry out the proof by induction on $\lambda(r)$; for $r = 1$ the statement of the lemma follows from (10). Let $r = ag$, where $a \in A_j$, $\lambda(g) = \lambda(r) - 1$ and therefore either $g = 1$ or else the first index k of g is different from j. If $j \neq i$, then by 3) of Lemma 1, $r_j(C) = r$, so that, again by 3) of Lemma 1, we have $r_j(gU) = g$. Hence, by the induction hypothesis

$$m(g, U) = 1. \tag{13}$$

Let $g \neq 1$. Then, by (9),

$$m(r, U) = m(ag, U) = m(a, gU) \cdot [gU, j, k] \cdot m(g, U).$$

However, it follows from $j \neq k$ (see the last paragraph of § 24.5) that $g = s(j, D)$, where $D = A_j g U$, so that, by (c) of § 24.7,

$$[gU, j, k] = 1.$$

From this and from (13) it follows that

$$m(r, U) = m(a, gU).$$

This equation obviously is also true for $g = 1$. Finally, since $a \in A_j$, we have, by the definition of the element $m(g, C)$,

$$m(r, U) = m(a, gU) = [r_j^{-1}(rU) \cdot a \cdot r_j(gU)]\sigma_{j,D}^{-1} = (r^{-1}ag)\sigma_{j,D}^{-1} = 1. \text{ //}$$

11. Lemma 7: *The product* $\sigma\tau$ *is the identity mapping of* B *onto itself.*

By (8), it is sufficient to find the images under $\sigma\tau$ of the elements of the form $[C, i, j]$ and the elements of the subgroups $B_{i,D}$.

If $b = [C, i, j]$, then

$$b(\sigma\tau) = (b\sigma_F)\tau = [r_i^{-1}(C) r_j(C)]\tau.$$

We set $r_i(C) = r$ and $r_j(C) = r_0$ and we denote the first indices of these elements by k and k_0, respectively. Then, from $r^{-1}r_0 \in U$ and (9) we have

$$b(\sigma\tau) = (r^{-1}r_0)\tau = m(r^{-1}r_0, U) = m(r^{-1}, r_0 U)\,[r_0 U, k, k_0]\, m(r_0, U).$$

But by Lemma 6 and the equation (11),

$$m(r_0, U) = 1,$$

$$m(r^{-1}, r_0 U) = m^{-1}(r, r^{-1}r_0 U) = m^{-1}(r, U) = 1;$$

and therefore, since $r_0 U = C$,

$$b(\sigma\tau) = [C, k, k_0].$$

If $k = i$, $k_0 = j$, then $b(\sigma\tau) = b$. If $k_0 \neq j$, then since $r_0 = r_j(C)$, we have $r_0 = s(j, D)$ for the corresponding D. Therefore, by (c) of § 24.7, $[r_0 U, j, k_0] = 1$, and then by (a) of § 24.7,

$$b(\sigma\tau) = [C, k, k_0] = [C, k, j]\,[C, j, k_0] = [C, k, j].$$

Finally, if in addition $k \neq i$, then a similar argument enables us to replace k by i, so that we again have $b(\sigma\tau) = b$.

Now let $b \in B_{i,D}$. Then

$$b\sigma = b\sigma_{i,D} = s^{-1}as, \tag{14}$$

where $a \in A_i$, $s = s(i, D)$. At the same time, $b\sigma \in U$, so that

$$asU = sU. \tag{15}$$

Hence by (9), if we let j denote the first index of s, $j \neq i$, we have

$$\begin{aligned} b(\sigma\tau) &= (s^{-1}as)\tau = m(s^{-1}as, U) = \\ &= m(s^{-1}, asU) \cdot [asU, j, i] \cdot m(a, sU) \cdot [sU, i, j] \cdot m(s, U). \end{aligned}$$

But by Lemma 6 and (11),

$$m(s, U) = 1,$$

$$m(s^{-1}, asU) = m^{-1}(s, s^{-1}asU) = m^{-1}(s, U) = 1.$$

Furthermore, by (c), (a), and (a') of § 24.7, and (15),

$$[sU, i, j] = 1,$$

$$[asU, j, i] = [sU, j, i] = [sU, i, j]^{-1} = 1.$$

Thus we see, by (14) and (15), that

$$b(\sigma\tau) = m(a, sU) = [r_i^{-1}(asU) \cdot a \cdot r_i(sU)]\,\sigma_{i,D}^{-1} = (s^{-1}as)\sigma_{i,D}^{-1} = b,$$

as required. //

12. From Lemmas 5 and 7 it follows that each of the mappings σ and τ is an isomorphism between U and B. And so, by (8), the definition of σ, and Lemma 3, the main statement of the Subgroup Theorem is proved. The second statement of this theorem is proved by the choice of representatives $s(i, D)$ made in § 24.5. //

Since every subgroup that is conjugate, in a certain group, to a subgroup of an infinite cyclic group is itself an infinite cyclic group, and since a free

product of free groups is itself a free group, the following theorem is an immediate consequence of the theorem just proved.

THE NIELSEN-SCHREIER THEOREM: *Every subgroup of a free group other than the unit subgroup is itself a free group.*

13. We note the following obvious property of free products:

If

$$G = \prod_{i \in I}{}^{*} A_i \tag{16}$$

and

$$A_i = \prod_{j \in J_i}{}^{*} A_{ij} \qquad (i \in I),$$

then

$$G = \prod_{i \in I}{}^{*} \prod_{j \in J_i}{}^{*} A_{ij}. \tag{17}$$

The free decomposition (17) is called a *refinement* of (16).

Two free decompositions of a group G,

$$G = \prod_{i \in I}{}^{*} A_i * F_1 = \prod_{j \in J}{}^{*} B_j * F_2,$$

are called *isomorphic* if F_1 and F_2 are isomorphic free groups and if a one-to-one correspondence between the free factors A_i and B_j can be established such that corresponding factors are conjugate in G.

The following theorem holds [A. G. Kurosh, *Math. Ann.*, vol. 109 (1934), pp. 647-60; R. Baer and F. Levi, *Comp. Math.*, vol. 3 (1936), pp. 391-98].

Any two free decompositions of an arbitrary group have isomorphic refinements.

For if

$$G = \prod_{i \in I}{}^{*} A_i = \prod_{j \in J}{}^{*} B_j, \tag{18}$$

then by applying the Subgroup Theorem, we obtain

$$B_j = \prod_{i \in I,\ D \in (B_j, A_i)}{}^{*} (B_j \cap s^{-1} A_i s) * F_j, \quad s = s(i, D) \qquad (j \in J);$$

$$A_i = \prod_{j \in J,\ D \in (B_j, A_i)}{}^{*} (A_i \cap t^{-1} B_j t) * F_i', \quad t = s(j, D') \qquad (i \in I).$$

Thus, we have determined refinements of the free decompositions (18), and we wish to prove that they are isomorphic.

We fix a pair of indices $i \in I$, $j \in J$ and note that if $D \in (A_i, B_j)$, then the set D^{-1} of the inverses of the elements of D is a class of the decomposition (B_j, A_i) and that in this way a one-to-one correspondence between the classes

of our two decompositions can be established. Let us show that the subgroups $B_j \cap s^{-1}A_i s$ and $A_i \cap t^{-1}B_j t$ are conjugate in G, where $s = s(i, D)$ and $t = s(j, D^{-1})$. Since

$$D^{-1} = B_j t A_i = (A_i s B_j)^{-1} = B_j s^{-1} A_i,$$

we have

$$t = b_j s^{-1} a_i \qquad (b_j \in B_j;\ a_i \in A_i).$$

Therefore

$$A_i \cap t^{-1}B_j t = A_i \cap a_i^{-1} s b_j^{-1} B_j b_j s^{-1} a_i = \\ = A_i \cap a_i^{-1} s B_j b_j s^{-1} a_i = a_i^{-1} s (s^{-1} A_i s \cap B_j) s^{-1} a_i.$$

To conclude the proof, it remains to point out that the normal subgroup generated in a group by a certain system of subgroups does not change if each of the subgroups is replaced by a conjugate, and then to make use of the followin property of a free product:

If $G = A * B$ and A generates the normal subgroup $\bar{A}$ in G, then B is isomorphic to the factor group $G/\bar{A}$.

For the normal subgroup $\bar{A}$ consists of precisely those elements of G for which the product of the elements of B that occur in their irreducible representation (with preservation of order) is equal to 1. Hence it follows that every coset of $\bar{A}$ in G contains exactly one element of B. //

CHAPTER IV

LATTICES

§ 25. Lattices and Complete Lattices

1. The parallelism between the theory of groups and the theory of rings has been responsible for the introduction into Algebra of other concepts in addition to that of an Ω-group. One of these concepts, that of a lattice, was suggested by the fact that certain sets—among them, the set of all subgroups, and the set of all normal subgroups of a group, the set of all subrings of a ring, the set of all ideals, and the set of all right ideals of a ring—are partially ordered by set-theoretical inclusion, where the partial ordering has certain additional properties.

A partially ordered set L is called a *lattice* if it satisfies the following two conditions:

1_1) For every pair of elements $a, b \in L$, there exists an element $c = a \cap b$ in L, the *intersection* of a and b, such that

$$c \leqq a, \qquad c \leqq b,$$

and if any element c' also has the properties $c' \leqq a$, $c' \leqq b$, then $c' \leqq c$.

1_2) For every pair of elements $a, b \in L$, there exists an element $d = a \cup b$ in L, the *union* of a and b, such that

$$d \geqq a, \qquad d \geqq b,$$

and if any element d' also has the properties $d' \geqq a$, $d' \geqq b$, then $d' \geqq d$.

It is clear that both the intersection $a \cap b$ and the union $a \cup b$ of two elements a and b are uniquely determined. It is equally clear that a partially ordered set inversely isomorphic to a lattice (see § 4.6) is itself a lattice, the concepts of intersection and union being dual to each other (see § 5.5).

Thus, we see that we can speak of the *lattice of subgroups* and of the *lattice of normal subgroups* of a certain group G, and also of the *lattice of subrings*, the *lattice of ideals*, the *lattice of left (or right) ideals* of a certain ring R. In all these cases the intersection of subgroups (or subrings) A

and B is their set-theoretical intersection $A \cap B$, and the role of the union is played by the subgroup (subring) $\{A, B\}$ generated by these subgroups (subrings). More generally, we can speak of the *lattice of subalgebras* of a given universal algebra and also of the *lattice of Ω-subgroups* and the *lattice of ideals* of a given Ω-group.

The following are a few further examples of lattices. All the subsets of a set M form a lattice under set-theoretical inclusion, where intersection and union have their set-theoretical meaning; this *lattice of subsets M* will be made use of in the sequel.

Every linearly ordered set L is a lattice, and if $a, b \in L$ and $a \leqq b$, then

$$a \cap b = a, \qquad a \cup b = b.$$

The set of natural numbers is a lattice if we take divisibility as the order relation. The role of intersection is played here by the greatest common divisor, and the union is the least common multiple.

2. Lattices can be regarded as a special case of universal algebras. For the concept of a lattice can be defined, without recourse to partial ordering, merely by means of the properties of the binary operations of intersection and union:

A set L with two binary operations $a \cap b$ and $a \cup b$ is a lattice if and only if these operations satisfy the following identical relations:

$2_1)$ $a \cap a = a$, $a \cup a = a$;

$2_2)$ $a \cap b = b \cap a$, $a \cup b = b \cup a$;

$2_3)$ $(a \cap b) \cap c = a \cap (b \cap c)$, $(a \cup b) \cup c = a \cup (b \cup c)$;

$2_4)$ $a \cap (a \cup b) = a$, $a \cup (a \cap b) = a$.

Let us assume first of all that L is a lattice, i.e., that the operations $a \cap b$ and $a \cup b$ are defined by the conditions $1_1)$ and $1_2)$. Then the properties $2_1)$ and $2_2)$ obviously hold. Let us verify $2_3)$, say, for intersection. Since, by $1_1)$,

$$(a \cap b) \cap c \leqq a \cap b \leqq a,$$

$$(a \cap b) \cap c \leqq a \cap b \leqq b,$$

$$(a \cap b) \cap c \leqq c,$$

we have, again by $1_1)$,

$$(a \cap b) \cap c \leqq b \cap c,$$

$$(a \cap b) \cap c \leqq a \cap (b \cap c).$$

Similarly,

$$a \cap (b \cap c) \leqq (a \cap b) \cap c,$$

so that $2_3)$ holds.

On the other hand, it is clear by 1_1) that

$$a \cap (a \cup b) \leqq a;$$

however, $a \leqq a$ and, by 1_2), $a \leqq a \cup b$, so that, by 1_1),

$$a \leqq a \cap (a \cup b).$$

Hence it follows that 2_4) holds.

Now let L be a set with two binary operations satisfying the properties 2_1)-2_4). If $a, b \in L$, then the equations

$$a \cap b = a, \qquad a \cup b = b \tag{1}$$

either both hold or both fail to hold. For if $a \cap b = a$, then by 2_4) and 2_2),

$$a \cup b = (a \cap b) \cup b = b;$$

and if $a \cup b = b$, then by 2_4),

$$a \cap b = a \cap (a \cup b) = a.$$

If the equations (1) hold for two elements a and b, then we set $a \leqq b$. Thus a partial ordering has been induced in L. For, by 2_1), $a \leqq a$. Further, if $a \leqq b$ and $b \leqq c$, i.e., $a \cap b = a$ and $b \cap c = b$, then by 2_3),

$$a \cap c = (a \cap b) \cap c = a \cap (b \cap c) = a \cap b = a,$$

i.e., $a \leqq c$. Finally, if $a \leqq b$ and $b \leqq a$, i.e., $a \cap b = a$ and $b \cap a = b$, then by 2_2), $a = b$.

Let us now show that the condition 1_1) is satisfied. It follows from $(a \cap b) \cap a = a \cap (a \cap b) = (a \cap a) \cap b = a \cap b$ that $a \cap b \leqq a$. Similarly, $a \cap b \leqq b$. And if c' is an arbitrary element of L satisfying the conditions $c' \leqq a$, $c' \leqq b$, i.e., $c' \cap a = c'$, $c' \cap b = c'$, then

$$c' \cap (a \cap b) = (c' \cap a) \cap b = c' \cap b = c',$$

and hence $c' \leqq a \cap b$. The element $a \cap b$ is therefore the intersection of a and b in the sense of condition 1_1). In a similar way, it can be shown that the element $a \cup b$ is the union of a and b in the sense of 1_2). //

3. Our second definition of a lattice shows that *lattices form a primitive class of universal algebras with two binary operations* (see § 22.3). The definitions, to be given below, of a sublattice and of isomorphic embedding of lattices indicate that it is this second definition that must be regarded essentially as the basic one.

In fact, a subset S of the lattice L is called a *sublattice* if it is a subalgebra of L regarded as a universal algebra in the sense of the above definition of § 25.2. In other words, when any two elements a and b belong to S, then so do their intersection $a \cap b$ and their union $a \cup b$ in the sense of these operations in L, i.e., S is itself a lattice with respect to the operations defined in L.

It should be noted that a subset S of a lattice L may well turn out to be a lattice under the partial ordering that is induced in S by the partial ordering given in L, without being a sublattice in the sense of the definition given above. For example, although the partial ordering in the lattice of subgroups of a group G is set-theoretical, this lattice is nevertheless not a sublattice in the lattice of all subsets of G, because unions have a different meaning in each of these two lattices. Similarly, the lattice of subrings of a ring R is not a sublattice of the lattice of subgroups of the additive group of the ring.

The lattice of normal subgroups of a group G is a sublattice of the lattice of all subgroups of the group, as was proved in § 13.5. More generally, from what was proved in § 18.4, it follows that *the lattice of ideals of an Ω-group G is a sublattice of the lattice of all Ω-subgroups of G, as well as of the lattice of subgroups of its additive group.*

4. A one-to-one mapping φ of a lattice L into a lattice L' is called an *isomorphic mapping* of L into L', or an *isomorphic embedding* of L in L', if for arbitrary $a, b \in L$

$$(a \cap b)\varphi = a\varphi \cap b\varphi, \qquad (a \cup b)\varphi = a\varphi \cup b\varphi.$$

In other words, it is an isomorphism of the lattices regarded as universal algebras (see § 17.5).

An isomorphic embedding φ of a lattice L in a lattice L' is also an isomorphic embedding of L in L' in the sense of isomorphism of partially ordered sets (see § 4.4).

For let $a, b \in L$. If $a \leqq b$, i.e., $a \cap b = a$, then

$$(a \cap b)\varphi = a\varphi \cap b\varphi = a\varphi,$$

and hence $a\varphi \leqq b\varphi$. By arguing in the opposite direction and using the fact that φ is one to one, we see that it follows from $a\varphi \leqq b\varphi$ that $a \leqq b$. //

Examples to be given below show that *the converse statement does not hold*: if L and L' are partially ordered sets and φ is an isomorphic mapping of L into L', then even when L and L' are lattices, φ need not be an isomorphic mapping of the lattice L into the lattice L'. But the situation is different when we consider isomorphic mappings of L *onto* L':

If L and L' are lattices, then every isomorphic mapping of L onto L' in the sense of partial ordering is also a lattice isomorphism of L onto L'.

For if $a, b \in L$, then from $a \cap b \leqq a$ it follows that $(a \cap b)\varphi \leqq a\varphi$; similarly, $(a \cap b)\varphi \leqq b\varphi$. If the element $c' \in L'$ is such that $c' \leqq a\varphi$ and $c' \leqq b\varphi$, and if c is that element of L for which $c\varphi = c'$, then $c \leqq a$ and $c \leqq b$, i.e., $c \leqq a \cap b$; hence $c\varphi \leqq (a \cap b)\varphi$. This shows that

$$(a \cap b)\varphi = a\varphi \cap b\varphi.$$

The proof for the union proceeds in the same way. //

We note that the concept of homomorphism could also be carried over to lattices and, quite generally, so could everything concerning arbitrary primitive classes of universal algebras; in particular, in accordance with § 23.2, the concept of a *free lattice* with a given set X of free generators is meaningful.

5. As follows from § 2.1, the binary relations of a set M form a lattice, which is the same as the lattice of all subsets of $M \times M$. Equivalence relations on M also form a lattice, as was shown in § 3.3. This *lattice of equivalence relations,* however, is not a sublattice of the lattice of binary relations.

❖ Every lattice can be isomorphically embedded in the lattice of equivalence relations defined on a suitable set [P. M. Whitman, *Bull. Amer. Math. Soc.*, vol. 52 (1946), pp. 507-22].

The lattice of equivalence relations defined on an arbitrarily given set can be isomorphically embedded in the lattice of subgroups of a suitable group [G. D. Birkhoff, *Proc. Camb. Phil. Soc.*, vol. 31 (1935) pp. 433-54].

Thus, every lattice can be isomorphically embedded in the lattice of subgroups of a suitable group. ❖

6. Many of the lattices mentioned above—for example, the lattice of subsets of a set M, the lattice of subgroups of a group G, the lattice of subrings of a ring R (or, more generally, the lattice of Ω-subgroups of an Ω-group G), and the lattice of equivalence relations on a set M—have the property that intersections and unions are defined in them not only for two elements, and hence, by associativity, for any finite number of elements, but also for arbitrary infinite subsets. In other words, these lattices are complete in the sense of the following definition:

A partially ordered set L is called a *complete lattice* if for every nonempty subset $A \subseteqq L$ of L there exist elements c and d with the following properties:

$1_1'$) For all $a \in A$ the inequality $c \leqq a$ holds, and if an element c' also satisfies the conditions $c' \leqq a$ for all $a \in A$, then $c' \leqq c$.

$1_2'$) For all $a \in A$ the inequality $d \geqq a$ holds, and if an element d' also satisfies the conditions $d' \geqq a$ for all $a \in A$, then $d' \geqq d$.

The uniquely determined elements c and d are called the *intersection* and the *union*, respectively, of the elements of A. They are written in the following way:

$$c = \bigcap_{a \in A} a, \qquad d = \bigcup_{a \in A} a$$

or, if the elements of A are denoted by a_α, where the index α ranges over a certain set M,

$$c = \bigcap_{\alpha \in M} a_\alpha, \qquad d = \bigcup_{\alpha \in M} a_\alpha.$$

Clearly, a complete lattice is also an ordinary lattice.

7. The intersection of all the elements of a complete lattice L is called the *null* of the lattice and is denoted by the symbol 0. This element is uniquely determined by any one of the following three conditions: For all $a \in L$,

$$1)\ 0 \leqq a; \quad 2)\ 0 \cap a = 0; \quad 3)\ 0 \cup a = a.$$

The union of all elements of a complete lattice is called the *unit* element, or the *all* element, of the lattice and is denoted by the symbol 1. This element is uniquely determined by any one of the following three conditions: For all $a \in L$,

$$1')\ 1 \geqq a; \quad 2')\ 1 \cup a = 1; \quad 3')\ 1 \cap a = a.$$

Lattices that are not complete may, of course, also have both a null and a unit element (or just one of them).

In the lattice of subgroups of a group G, the roles of the null and of the unit element are played, respectively, by the unit subgroup and the group G itself; in the lattice of subrings of a ring R, they are the zero subring and the ring R itself; in the lattice of subsets of a set M, they are the empty subset and M itself. In the lattice of natural numbers ordered by divisibility and in the chain of natural numbers in their natural order the null is the number 1, but there is no unit element.

8. *A partially ordered set L is a complete lattice if and only if it has a unit element and any collection of its non-empty subsets have an intersection.*

We need only prove the statement that if a unit element and all the intersections exist in L, then the unions also exist. Let A be a non-empty subset of L. L contains elements b such that $b \geqq a$ for all $a \in A$; 1 is such

an element in any case. Let B be the non-empty set of all these elements b, and d their intersection:

$$d = \bigcap_{b \in B} b.$$

We shall show that d is the union of the elements of A. For if $a \in A$, then $a \leqq b$ for all $b \in B$, and therefore $a \leqq d$. On the other hand, if an element $l \in L$ is such that $l \geqq a$ for all $a \in A$, then $l \in B$, and therefore $d \leqq l$. Thus,

$$d = \bigcup_{a \in A} a$$

holds. //

Since the property of being a complete lattice (like the property of being a lattice) is preserved under an inverse isomorphism of partially ordered sets (see § 4.6), in the theorem just proved we can also require the existence of a null and of the unions of all non-empty subsets.

❖ Let a mapping φ of a partially ordered set M into a partially ordered set N be called *monotonic* if for $a, b \in M$ it always follows from $a \leqq b$ that $a\varphi \leqq b\varphi$. A lattice L is complete if and only if for every monotonic mapping φ of L into itself there exist *fixed* elements, that is, elements a such that $a\varphi = a$ [A. Tarski, *Pacif. J. Math.*, vol. 5 (1955), pp. 285-309; A. C. Davis, *Pacif. J. Math.*, vol. 5 (1955), pp. 311-19]. ❖

9. It was shown in § 4.5 that every partially ordered set M can be embedded isomorphically in the complete lattice $\tilde{M}$ of all its subsets. We now wish to prove the following theorem:

Every lattice can be embedded isomorphically, in the sense of lattice isomorphism (see § 25.4), *in a complete lattice.*

This theorem is a consequence of the following statement, which is a stronger form of the theorem of § 4.5:

Every partially ordered set M can be embedded isomorphically in a complete lattice L in such a way that if the intersection or the union of a subset $A \subseteq M$ exists in M, then it is preserved in this embedding φ:

$$\Big(\bigcap_{a \in A} a\Big)\varphi = \bigcap_{a \in A} a\varphi, \quad \Big(\bigcup_{a \in A} a\Big)\varphi = \bigcup_{a \in A} a\varphi. \tag{2}$$

It is clear that the embedding of M into $\tilde{M}$ that was constructed in § 4.5 fails to satisfy these requirements. We proceed as follows. First of all, let us assume that the set M has a null 0, for otherwise we could simply add it to M. A non-empty subset $X \subseteq M$ will be called an *ideal* in M if the following conditions are satisfied: 1) If X contains the element x, then it also contains all $y \in M$ for which $y \leqq x$; 2) For every $X' \subseteq X$ for which a union exists in M, this union is contained in X.

The set L of all ideals of M is partially ordered by set-theoretical inclusion. Since L has a unit element—M itself is, of course, an ideal—and since the set-theoretical intersection of an arbitrary set of ideals contains a null and satisfies all the conditions in the definition of an ideal, L is a complete lattice, by § 25.8.

For each element $a \in M$ the set (a) of all elements $b \in M$ with $b \leqq a$ is an ideal; this is the *principal ideal* generated by a. By assigning to every $a \in M$ the principal ideal (a) generated by a, we obtain, as in § 4.5, an isomorphic embedding q of the partially ordered set M in the complete lattice L.

Let us now assume that for a subset $A \subseteq M$ of M there exists the intersection

$$c = \bigcap_{a \in A} a.$$

Then $c \leqq a$, and therefore $(c) \subseteq (a)$ for all $a \in A$. If the set-theoretical intersection of all principal ideals (a) $(a \in A)$ contains an element x, then $x \leqq a$ for all $a \in A$, and therefore $x \leqq c$. Thus,

$$(c) = \bigcap_{a \in A} (a),$$

and this proves the first of the equations (2).

Finally, suppose that for a subset $A \subseteq M$ of M there exists the union

$$d = \bigcup_{a \in A} a.$$

Then $(d) \supseteq (a)$ for all $a \in A$. If a certain ideal X contains all the elements $a \in A$, then by the definition of an ideal, it also contains the union d, and hence $(d) \subseteq X$. Thus,

$$(d) = \bigcup_{a \in A} (a).$$

This proves the second of the equations (2). //

The reader will have no difficulty in verifying that the construction used in this proof, when carried out on the ordered set of rational numbers, in essence gives the usual completion of the system of rational numbers by Dedekind cuts.

10. *A partially ordered set L is a complete lattice if and only if L has a null and every ideal in L is a principal ideal.*

For if L is a complete lattice, then for every ideal X the union a of all its elements exists and, by the definition of an ideal, $a \in X$, so that $X = (a)$. Conversely, suppose that a partially ordered set L contains a null and every

ideal in L is a principal ideal. In particular, L itself is a principal ideal and therefore has a unit element. If A is a non-empty subset of L, then the collection X of all those elements $x \in L$ for which $x \leqq a$ for all $a \in A$ contains 0 and is an ideal. This ideal must be principal, $X = (b)$, and b is the intersection of the elements of A. By § 25.8, L is a complete lattice. //

Thus, the construction given in the preceding subsection, when carried out on a complete lattice, does not yield anything new.

§ 26. Modular Lattices

1. Of the axioms 2_1)-2_4) defining a lattice (see § 25.2), only 2_4) connects the operations of intersection and union. This connection is very weak. Certain lattices—for example, the lattice of all subsets of a set—satisfy the additional condition of distributivity:

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c) \tag{1}$$

(see § 1.1). Lattices satisfying this condition for arbitrary a, b, and c are called *distributive* lattices and will be studied in § 40. However, the lattices of normal subgroups and, a fortiori, the lattices of all subgroups are not, as a rule, distributive.

With the intention of singling out, if possible, a restricted class of lattices comprising the lattices of normal subgroups of arbitrary groups (though not necessarily the lattices of all subgroups) we make the following definition:

A lattice L is called a *modular* (or a *Dedekind*) lattice if the equation

$$a \cap (b \cup c) = b \cup (a \cap c) \tag{2}$$

holds for any $a, b, c \in L$ for which $a \geqq b$.

Every distributive lattice is modular, because it follows from $a \geqq b$ that $a \cap b = b$ and hence, in virtue of $a \geqq b$, (1) implies (2).

Every sublattice S of a modular lattice L (see § 25.3) *is itself modular,* because (2) must hold, in particular, for elements $a, b, c \in S$ satisfying the condition $a \geqq b$.

It is quite obvious that the property of being modular is preserved under lattice isomorphism. Also, *a lattice inversely isomorphic to a modular lattice is itself modular,* because changing the symbol $\geqq$ in the definition of a modular lattice into $\leqq$ and interchanging the symbols $\cap$ and $\cup$ results in the same definition, but with the letters a and b interchanged.

2. *The lattice of all normal subgroups of an arbitrary group is modular.*

For let A, B, and C be normal subgroups of a group G, with $A \supseteq B$. We have to show, taking (1) of § 20.1 into account, that

$$A \cap BC = B(A \cap C). \tag{3}$$

Since $B \subseteq A$ and $B \subseteq BC$, B is contained in the left-hand side of (3). This is true of $A \cap C$ also, because $C \subseteq BC$. This implies the inclusion

$$A \cap BC \supseteq B(A \cap C). \tag{4}$$

On the other hand, every element contained in the normal subgroup $A \cap BC$ is an element $a \in A$ that at the same time admits an expression

$$a = bc, \tag{5}$$

where $b \in B$, $c \in C$. Hence $c = b^{-1}a \in A$, because $B \subseteq A$, i.e.,

$$c \in (A \cap C).$$

Therefore, by (5), every element of the left-hand side of (3) is also contained in the right-hand side, and this, together with (4), proves their equality. //

By what was proved in § 18.4, *the ideals of an arbitrary Ω-group form a sublattice of the lattice of normal subgroups of the additive group of the Ω-group, and thus they constitute a modular lattice.* Hence it follows that *the lattice of all ideals of an arbitrary ring is modular.*

Note that the lattice of all subgroups of a non-commutative group, as a rule, is not modular.

3. There are many different definitions of a modular lattice equivalent to the one given above. Let us indicate a few of them.

A lattice L is modular if and only if for arbitrary $a, b, c \in L$

$$a \cap [(a \cap b) \cup c] = (a \cap b) \cup (a \cap c). \tag{6}$$

For if a lattice L is modular, then (6) follows from (2), because $a \geqq a \cap b$. Conversely, if (6) holds in a lattice L, then for $a \geqq b$ (2) follows from (6), because in this case $a \cap b = b$. //

Thus we see that *modular lattices form a primitive class of universal algebras.*

4. *A lattice L is modular if and only if the conditions*

$$a \leqq b, \quad a \cap c = b \cap c, \quad a \cup c = b \cup c \tag{7}$$

for elements $a, b, c \in L$ imply that $a = b$.

For if a lattice L is modular and the elements a, b, c satisfy the conditions (7), then

$$a = a \cup (a \cap c) = a \cup (b \cap c) = b \cap (a \cup c) = b \cap (b \cup c) = b.$$

Now suppose that the premises of the converse statement are satisfied and that a, b, c are elements of L with $a \geqq b$. It is easy to verify (see the derivation of (4)) that no matter what lattice L is, it follows from $a \geqq b$ that

$$a \cap (b \cup c) \geqq b \cup (a \cap c). \tag{8}$$

Furthermore,

$$[a \cap (b \cup c)] \cap c = a \cap [(b \cup c) \cap c] = a \cap c, \tag{9}$$

and since for $a \geqq b$

$$a \geqq b \cup (a \cap c),$$

we have

$$a \cap c \geqq [b \cup (a \cap c)] \cap c \geqq (a \cap c) \cap c = a \cap c,$$

and thus

$$[b \cup (a \cap c)] \cap c = a \cap c. \tag{10}$$

From (9) and (10) it follows that for $a \geqq b$

$$[a \cap (b \cup c)] \cap c = [b \cup (a \cap c)] \cap c. \tag{11}$$

By going over from (11) to the dual equation—i.e., by interchanging the symbols $\cap$ and $\cup$—and, in addition interchanging the letters a and b, we obtain

$$[a \cap (b \cup c)] \cup c = [b \cup (a \cap c)] \cup c, \tag{12}$$

where the inequality $a \geqq b$ is preserved, because the symbol $\geqq$ must also be replaced by $\leqq$. Thus we see that (8), (11), and (12) have the form of the conditions (7) and therefore, by our assumptions, the equation (2) holds for $a \geqq b$. //

5. The properties of invariant and principal series of Ω-groups, which we are familiar with from § 20.7, could be carried over to the theory of modular lattices. We shall not do this here to the fullest extent and shall confine ourselves to results required later.

Let L be a modular lattice having a null and a unit element. An ordered finite system of elements,

$$0 = a_0 < a_1 < a_2 < \ldots < a_{k-1} < a_k = 1 \tag{13}$$

is called a *normal series* of the lattice;[1] the number k is the *length* of the series. A normal series

$$0 = b_0 < b_1 < b_2 < \ldots < b_{l-1} < b_l = 1 \tag{14}$$

is called a *refinement* of (13) if every element a_i $(i = 0, 1, \ldots, k)$ is equal to one of the elements b_j $(0 \leqq j \leqq l)$.

6. *Any two normal series of a modular lattice L have refinements of equal length.*

For let (13) and (14) be arbitrary normal series of L. By repeating to a certain degree what we have done in § 20.5, we set

$$a_{ij} = a_i \cup (a_{i+1} \cap b_j) \qquad (i = 0, 1, \ldots, k-1;\ j = 0, 1, \ldots, l),$$
$$b_{ji} = b_j \cup (b_{j+1} \cap a_i) \qquad (j = 0, 1, \ldots, l-1;\ i = 0, 1, \ldots, k).$$

Since

$$a_{i0} = a_i, \qquad a_{il} = a_{i+1},$$

and

$$a_{ij} \leqq a_{i,j+1} \qquad (j = 0, 1, \ldots, l-1),$$

the elements a_{ij} form a normal series, with possible repetitions, that is a refinement of (13). Similarly, the elements b_{ji} form a refinement of (14). These two new series have the same length kl, so that it now remains to prove that they have the same number of repetitions.

Suppose, then, that

$$a_{ij} = a_{i,j+1}. \tag{15}$$

Using, in succession, the definition of the element $a_{i,j+1}$, the equation (15), the inequality $a_{ij} \leqq a_{i+1}$, the definition of the element a_{ij}, and finally, the definition (2) of a modular lattice together with the inequality $a_{i+1} \cap b_j \leqq b_{j+1}$, we obtain:

$$\begin{aligned} a_{i+1} \cap b_{j+1} &= a_{i,j+1} \cap (a_{i+1} \cap b_{j+1}) = a_{ij} \cap (a_{i+1} \cap b_{j+1}) = \\ &= a_{ij} \cap b_{j+1} = [a_i \cup (a_{i+1} \cap b_j)] \cap b_{j+1} = \\ &= (a_i \cap b_{j+1}) \cup (a_{i+1} \cap b_j). \end{aligned}$$

This result, in conjunction with the definition of the elements b_{ji} and $b_{j,i+1}$ and the obvious inequality $a_{i+1} \cap b_j \leqq b_j$ leads to the equations

$$\begin{aligned} b_{ji} &= b_j \cup (b_{j+1} \cap a_i) = b_j \cup (b_{j+1} \cap a_i) \cup (a_{i+1} \cap b_j) = \\ &= b_j \cup (a_{i+1} \cap b_{j+1}) = b_{j,i+1}. \end{aligned}$$

[1] This is the accepted terminology, although for consistency we should speak of invariant series.

Thus, we have shown that in the refinements constructed above of the normal series (13) and (14), the repetitions are in one-to-one correspondence, so that they can be omitted simultaneously. //

7. We define a *principal series* of a modular lattice as a normal series that cannot be refined without repetitions. As in § 20.6, the theorem proved above has the following consequences:

If a modular lattice has principal series, then all its principal series are of equal length.

If a modular lattice has principal series, then each of its normal series can be refined to a principal series.

From this it follows that *if a modular lattice L has principal series, then every sublattice S of L also has principal series.* For if the principal series of L have the length k, then the length of a chain of the sublattice S cannot exceed k. If

$$c_0 < c_1 < \ldots < c_l \tag{16}$$

is one of the chains of maximal length contained in S, then c_0 and c_l are the null and the unit element, respectively, of S, and (16) is one of its principal series. //

8. These results enable us to state yet another form of the definition of a modular lattice:

A lattice L is modular if and only if in every sublattice having principal series all the principal series are of equal length.

For if a lattice L is modular, then it is sufficient to refer to § 26.1 and § 26.7. But if L is not modular, then it follows from § 26.4 that L contains elements a, b, c for which the condition (7) holds, although $a \neq b$. It is easy to verify that under these conditions the elements $a, b, c, a \cap c$, and $a \cup c$ form a sublattice of L and that the last two elements are the null and the unit element, respectively, of this sublattice. However, this sublattice has two principal series of distinct lengths, namely,

$$a \cap c < a < b < a \cup c,$$
$$a \cap c < c < a \cup c. //$$

9. *A modular lattice L has principal series if and only if it satisfies the descending and ascending chain conditions* (see §§ 5.1 and 5.5).

For if a modular lattice L has principal series of length k, then the lengths of its normal series do not exceed k. However, every finite chain of L, after the addition, if necessary, of the null and the unit element, is a normal series of the lattice. This shows that L cannot have infinite chains.

Conversely, suppose that in a lattice L the descending and ascending chain conditions, and consequently the conditions equivalent to them (see § 5.1), hold. In virtue of the minimal condition, there exist minimal elements of L; and since the intersection of two minimal elements must coincide with each of them, L actually contains only one such element; it is the null of our lattice. Similarly, L also contains a unit element.

If $a \in L$ and $a \neq 1$, then the non-empty set of elements x such that $x > a$ contains minimal elements—or, as we sometimes say, elements that *cover* a. Using the Axiom of Choice, for every a $(a \neq 1)$ we can select one of its covering elements; we denote this covering element by a'. Then the chain

$$0 = a_0 < a_1 < a_2 < \ldots < a_n < \ldots, \tag{17}$$

where

$$a_n = a'_{n-1} \qquad (n = 1, 2, \ldots),$$

must terminate, by hypothesis. Consequently, there exists an n such that $a_n = 1$, that is, the chain (17) is a principal series of L. //

10. If, in an arbitrary lattice L, elements a and b with $a \geqq b$ are chosen, then the collection of elements x such that $a \geqq x \geqq b$ is a sublattice of L having b as its null and a as its unit element. We denote this sublattice by a/b.

If arbitrary elements a and b are chosen in a modular lattice L, then the sublattices $(a \cup b)/a$ and $b/(a \cap b)$ are isomorphic.

For if $x \in b/(a \cap b)$, i.e.,

$$a \cap b \leqq x \leqq b,$$

then we set

$$x\varphi = x \cup a.$$

It is clear that $a \leqq x\varphi \leqq a \cup b$, or $x\varphi \in (a \cup b)/a$.

If there were an element y $(a \cap b \leqq y \leqq b)$, distinct from x, such that $y\varphi = x\varphi$, then the element $z = x \cup y$ would necessarily be distinct from at least one of the elements x and y, say from the first: $z > x$. However,

$$a \cap b \leqq z \leqq b,$$

so that in virtue of $a \cup a = a$ we have

$$z \cup a = x \cup y \cup a = (x \cup a) \cup (y \cup a) = x\varphi \cup y\varphi = x\varphi = x \cup a,$$

$$z \cap a = a \cap b = x \cap a.$$

By § 26.4, this contradicts the fact that L is modular.

On the other hand, if $x' \in (a \cup b)/a$, i.e.

$$a \leqq x' \leqq a \cup b,$$

then, as the lattice is modular,

$$x' = (a \cup b) \cap x' = a \cup (b \cap x'),$$

and since

$$a \cap b \leqq b \cap x' \leqq b,$$

we have

$$x' = (b \cap x')\varphi.$$

Thus, φ proves to be a one-to-one mapping of the sublattice $b/(a \cap b)$ onto the whole sublattice $(a \cup b)/a$. Let us show that this is an isomorphism. If $x, y \in b/(a \cap b)$, then it follows from $x \leqq y$ that $x \cup a \leqq y \cup a$, i.e., $x\varphi \leqq y\varphi$. On the other hand, if $x\varphi \leqq y\varphi$, i.e., $x \cup a \leqq y \cup a$, then

$$(x \cup y) \cup a = (x \cup a) \cup (y \cup a) = y \cup a,$$

i.e., $(x \cup y)\varphi = y\varphi$, and as φ is a one-to-one mapping, this implies that $x \cup y = y$, i.e., $x \leqq y$. It now remains to refer to the last result of § 25.4. //

11. Suppose that a modular lattice L has principal series. Then for every $a \in L$ the sublattice $a/0$ also has principal series; we denote the length of these series by $l(a)$.

For arbitrary elements a and b of a modular lattice L having principal series the following equation holds:

$$l(a \cup b) = l(a) + l(b) - l(a \cap b). \tag{18}$$

For it has been shown above that the sublattices $(a \cup b)/a$ and $b/(a \cap b)$ are isomorphic. The principal series of these two sublattices therefore have the same length. However, for the first sublattice, this length is $l(a \cup b) - l(a)$ and for the second sublattice, it is $l(b) - l(a \cap b)$. //

§ 27. Direct Unions. The Schmidt-Ore Theorem

1. A frequently used construction in the theory of groups is the so-called *direct product* of groups or, when the group operation is written additively, their *direct sum*. The corresponding role in the theory of rings is played by the *two-sided direct sum* of rings. The study of these concepts leads to completely parallel theories, which are in fact special manifestations of a theory concerning modular lattices. We shall conduct the discussion on

this level of generality, supplementing it in the next section with observations pertaining especially to groups, rings, and Ω-groups.

2. Let L be a modular lattice having a null. An element $a \in L$ is called the *direct union* of elements $b_1, b_2, \ldots, b_k \in L$ and is written in the form

$$a = b_1 \times b_2 \times \ldots \times b_k \tag{1}$$

if a is the union of these elements,

$$a = b_1 \cup b_2 \cup \ldots \cup b_k \tag{2}$$

and if for $i = 1, 2, \ldots, k$

$$(b_1 \cup \ldots \cup b_{i-1} \cup b_{i+1} \cup \ldots \cup b_k) \cap b_i = 0. \tag{3}$$

If L is a complete lattice (see § 25.6), then we can also define infinite direct unions: an element a is the *direct union* of elements b_i, where i ranges over an index set I, if

$$a = \bigcup_{i \in I} b_i \tag{4}$$

and if for all $i \in I$,

$$b_i \cap \bar{b}_i = 0, \tag{5}$$

where

$$\bar{b}_i = \bigcup_{j \in I,\ j \neq i} b_j.$$

3. Let us begin with an investigation of the basic properties of direct unions in a modular lattice L with a null. We confine ourselves here to the case of finite direct unions. The fact that L is modular, which played no role in the definitions of § 27.2, will now be put to essential use in many places.

In the definitions of § 27.2, *condition* (3) *can be replaced by the following equivalent condition*:

$$(b_1 \cup \ldots \cup b_{i-1}) \cap b_i = 0 \qquad (i = 2, 3, \ldots, k). \tag{6}$$

It is clear that (6) follows from (3). To prove that (3) follows from (6) we shall show by induction on l that the intersection of each of the l elements $b_1, b_2, \ldots, b_l$ $(2 \leqq l \leqq k)$ with the union of the remaining $l - 1$ elements of the system is equal to the null. Indeed, for $l = 2$ this statement follows from (6) for $i = 2$. Suppose that our statement is already proved for $l - 1$. From (6), we have the equation

$$(b_1 \cup \ldots \cup b_{l-1}) \cap b_l = 0. \tag{7}$$

And if $j < l$, then by employing the inequality

$$b_j \leqq b_1 \cup \ldots \cup b_{l-1},$$

the fact that L is modular, the equation (7), and the induction hypothesis, we obtain

$$\begin{aligned}(b_1 \cup \ldots \cup b_{j-1} \cup b_{j+1} \cup \ldots \cup b_l) \cap b_j = \\ = [(b_1 \cup \ldots \cup b_{j-1} \cup b_{j+1} \cup \ldots \cup b_{l-1}) \cup b_l] \cap (b_1 \cup \ldots \cup b_{l-1}) \cap b_j = \\ = (b_1 \cup \ldots \cup b_{j-1} \cup b_{j+1} \cup \ldots \cup b_{l-1}) \cap b_j = 0.\end{aligned}$$

Our statement is therefore proved for all l, and in particular, for $l = k$, and this leads to the condition (3). //

4. With the help of (6) the reader can immediately prove the following properties of a direct union:

If $a = b_1 \times b_2 \times \ldots \times b_k$ *and* $b_1 \cup b_2 \cup \ldots \cup b_l = c$ $(l < k)$, *then*

$$a = c \times b_{l+1} \times \ldots \times b_k.$$

If $a = b_1 \times b_2 \times \ldots \times b_k$ *and* $b_l = c_1 \times c_2 \times \ldots \times c_l$, *then*

$$a = c_1 \times c_2 \times \ldots \times c_l \times b_2 \times \ldots \times b_k.$$

Bearing in mind that in the definitions of § 27.2 the numbering of the elements $b_1, b_2, \ldots, b_k$ is immaterial, we arrive at the following result by repeated application of the statements just proved:

If

$$a = b_1 \times b_2 \times \ldots \times b_k \tag{8}$$

and if some or all of the elements b_i *are themselves decomposed into direct unions,*

$$b_i = c_{i1} \times c_{i2} \times \ldots \times c_{il_i} \qquad (1 \leqq l_i;\ i = 1, 2, \ldots, k),$$

then a *is the direct union of all the elements* c_{ij} $(1 \leqq j \leqq l_i;\ i = 1, 2, \ldots, k)$. This new direct decomposition of a is called a *refinement* of the direct decomposition (8).

On the other hand, it follows immediately from the definition of a direct union that *if we choose elements* $b_1, b_2, \ldots, b_l$ $(l < k)$ *in* (8) *and set* $b_1 \cup b_2 \cup \ldots \cup b_l = c$, *then*

$$c = b_1 \times b_2 \times \ldots \times b_l.$$

From this, and from the first statement of the present subsection, we derive the following result:

If (8) *is a direct decomposition of an element a and if the system of elements* $b_1, b_2, \ldots, b_k$ *is split into disjoint subsystems, then on replacing each of the subsystems by the union of its elements, we obtain a new direct decomposition of a for which* (8) *is a refinement.*

5. *Let* (8) *be a direct decomposition of an element a. If elements* b_i' *are chosen,* $0 \leqq b_i' \leqq b_i$ $(i = 1, 2, \ldots, k)$ *and if*

$$a' = b_1' \cup b_2' \cup \ldots \cup b_k',$$

then

$$a' = b_1' \times b_2' \times \ldots \times b_k'. \tag{9}$$

If $b_i' < b_i$ *for at least one i, then* $a' < a$.

In fact, for $i = 1, 2, \ldots, k$, we have

$$(b_1' \cup \ldots \cup b_{i-1}' \cup b_{i+1}' \cup \ldots \cup b_k') \cap b_i' \leqq$$
$$\leqq (b_1 \cup \ldots \cup b_{i-1} \cup b_{i+1} \cup \ldots \cup b_k) \cap b_i = 0,$$

which proves (9). On the other hand, if $a' = a$, then, making use of the fact that L is a modular lattice, we obtain for $i = 1, 2, \ldots, k$:

$$b_i = b_i \cap a = b_i \cap a' = b_i \cap (b_1' \cup b_2' \cup \ldots \cup b_k') =$$
$$= b_i' \cup [b_i \cap (b_1' \cup \ldots \cup b_{i-1}' \cup b_{i+1}' \cup \ldots \cup b_k')] = b_i',$$

because

$$b_i \cap (b_1' \cup \ldots \cup b_{i-1}' \cup b_{i+1}' \cup \ldots \cup b_k') \leqq$$
$$\leqq b_i \cap (b_1 \cup \ldots \cup b_{i-1} \cup b_{i+1} \cup \ldots \cup b_k) = 0. \text{ //}$$

6. *If* $a = b \times c$ *and the element* a' *satisfies the conditions* $b \leqq a' \leqq a$, *then*

$$a' = b \times (a' \cap c).$$

For, in virtue of the modularity,

$$a' = a' \cap a = a' \cap (b \cup c) = b \cup (a' \cap c).$$

On the other hand,

$$b \cap (a' \cap c) \leqq b \cap c = 0. \text{ //}$$

7. We note that if a modular lattice L has principal series and if $a = b \times c$, then by (18) of § 26.11,

$$l(a) = l(b) + l(c). \tag{10}$$

More generally, if $a = b_1 \times b_2 \times \ldots \times b_k$, then

$$l(a) = l(b_1) + l(b_2) + \ldots + l(b_k). \tag{11}$$

8. As a rule, whenever use is made of direct unions, the isomorphism problem for distinct direct decompositions of a given group (or a given ring) plays an important role. This problem has been the object of numerous investigations, primarily within the framework of the theory of modular lattices. Here we restrict our account to one special, but important result that was the starting point of these investigations.

We introduce one auxiliary concept. Suppose that L is a modular lattice with a null 0 and a unit element 1, and that a direct decomposition is given:

$$1 = a_1 \times a_2 \times \ldots \times a_k. \tag{12}$$

We set, as usual,

$$\bar{a}_i = a_1 \cup \ldots \cup a_{i-1} \cup a_{i+1} \cup \ldots \cup a_k \qquad (i = 1, 2, \ldots, k). \tag{13}$$

Hence, by § 27.4,

$$1 = a_i \times \bar{a}_i \qquad (i = 1, 2, \ldots, k). \tag{14}$$

If $b \in L$, then we define the *component* of this element in the factor a_i of the direct decomposition (12) as the element

$$b^i = a_i \cap (b \cup \bar{a}_i). \tag{15}$$

9. *The element b is contained in the union of all its components relative to the given direct decomposition* (12),

$$b \leqq b^1 \cup b^2 \cup \ldots \cup b^k.$$

For, using (15) and applying the definition of a modular lattice several times, we obtain:

$$\begin{aligned}
&b^1 \cup b^2 \cup \ldots \cup b^k = [a_1 \cap (b \cup \bar{a}_1)] \cup [a_2 \cap (b \cup \bar{a}_2)] \cup \ldots \cup [a_k \cap (b \cup \bar{a}_k)] = \\
&= \{a_1 \cup [a_2 \cap (b \cup \bar{a}_2)] \cup \ldots \cup [a_k \cap (b \cup \bar{a}_k)]\} \cap (b \cup \bar{a}_1) = \\
&= \{a_1 \cup a_2 \cup [a_3 \cap (b \cup \bar{a}_3)] \cup \ldots \cup [a_k \cap (b \cup \bar{a}_k)]\} \cap (b \cup \bar{a}_1) \cap (b \cup \bar{a}_2) = \\
&= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots = \\
&= (a_1 \cup a_2 \cup \ldots \cup a_k) \cap (b \cup \bar{a}_1) \cap (b \cup \bar{a}_2) \cap \ldots \cap (b \cup \bar{a}_k) = \\
&= (b \cup \bar{a}_1) \cap (b \cup \bar{a}_2) \cap \ldots \cap (b \cup \bar{a}_k) \geqq b. \;//
\end{aligned}$$

Note that in this proof we have only used inequalities such as, for example,

$$a_i \cap (b \cup \bar{a}_i) \leqq a_i \leqq b \cup \bar{a}_1 \qquad (i \neq 1).$$

10. We call two elements a and b of our lattice L *directly similar* if there exists an element c in L such that

$$1 = a \times c = b \times c.$$

We call two direct decompositions of the unit element of a lattice L *directly similar* if they consist of an equal number of factors and if a one-to-one correspondence can be set up between these direct factors such that corresponding factors are directly similar in L.

Finally, we call a non-null element a of a lattice L *indecomposable* if it cannot be decomposed into the direct union of two elements both distinct from the null.

11. If

$$1 = a_1 \times a_2 \times \ldots \times a_k$$

is a direct decomposition of the unit element of the lattice L, then by setting

$$c_l = a_1 \times a_2 \times \ldots \times a_l \qquad (l \leqq k),$$

we obtain a normal series of length k in L:

$$0 = c_0 < c_1 < c_2 < \ldots < c_k = 1.$$

Hence it follows, by § 26.7, that *if a modular lattice L has principal series and if the length of these series is n, then each direct decomposition of the unit element consists of not more than n factors and can therefore be refined* (see § 27.4) *to a direct decomposition with indecomposable factors.*

12. Our goal is a proof of the following theorem [O. J. Schmidt, *Math. Zeitschr.*, vol. 29 (1928), pp. 34-41; O. Ore, *Ann. of Math.*, vol. 37 (1936), pp. 265-92].

In a modular lattice L having principal series, any two decompositions of the unit element into the direct union of indecomposable elements are directly similar.

This theorem is a consequence of the following theorem:

If

$$1 = a_1 \times a_2 \times \ldots \times a_k, \tag{16}$$

$$1 = b_1 \times b_2 \times \ldots \times b_l \tag{17}$$

are any two decompositions of the unit element of a modular lattice L having principal series into the direct union of indecomposable elements, then every factor of each of these decompositions can be replaced by a suitable factor of the other decomposition.

Here the possibility of *replacing* the element a_1 in (16) by a factor b_j of (17) must be interpreted as the existence of a direct decomposition

$$1 = b_j \times a_2 \times \ldots \times a_k,$$

that is, by (14),

$$1 = b_j \times a_1. \tag{18}$$

The first theorem above follows from the second. For if the element b_j of (17) replaces the element a_1 in (16), then the equations (14) (for $i = 1$) and (18) show that the elements a_1 and b_j are directly similar in L. Suppose now that we have already constructed the direct decomposition

$$1 = b_{j_1} \times \ldots \times b_{j_m} \times a_{m+1} \times \ldots \times a_k \quad (1 \leqq m < k), \tag{19}$$

where $b_{j_1}, \ldots, b_{j_m}$ are distinct factors of (17) and the elements a_i and b_{j_i} ($i = 1, 2, \ldots, m$) are directly similar in L. We apply the second theorem to the direct decompositions (19) and (17). The element a_{m+1} in (19) must be replaceable by a factor $b_{j_{m+1}}$ of (17), so that a_{m+1} and $b_{j_{m+1}}$ are directly similar. Here the index j_{m+1} is distinct from all the indices $j_1, \ldots, j_m$, because from

$$1 = b_{j_1} \times \ldots \times b_{j_m} \times b_{j_{m+1}} \times a_{m+2} \times \ldots \times a_k$$

it follows that

$$b_{j_1} \cap b_{j_{m+1}} = 0 \qquad (i = 1, 2, \ldots, m).$$

Continuing in this way we finally arrive at a direct decomposition

$$1 = b_{j_1} \times b_{j_2} \times \ldots \times b_{j_k},$$

and hence $k \leqq l$. Comparing this with (17), we see that in fact $k = l$ and that the direct decompositions (16) and (17) are directly similar. //

13. We shall prove the second theorem by induction on the length of the principal series of the lattices in question; if the principal series is of length 1, the lattice consists of two elements—the null and the unit element—and for this lattice the theorem to be proved is obviously true.

Now suppose that the element a_1 of the first decomposition (16) is to be replaced. We denote by a_1^i the component (see § 27.8) of a_1 in the direct factor b_i of (17). To begin with, we shall assume that for at least one i the element a_1^i is different from b_i. When we now set

$$g = a_1^1 \cup a_1^2 \cup \ldots \cup a_1^l,$$

then we obtain, by § 27.5,

$$g = a_1^1 \times a_1^2 \times \ldots \times a_1^l, \tag{20}$$

with $g < 1$. Since, by § 27.9, $a_1 \leqq g$, we have, by § 27.6,

$$g = a_1 \times d, \tag{21}$$

where

$$d = g \cap (a_2 \times \ldots \times a_k). \tag{22}$$

By the induction hypothesis, the theorem is already proved for the lattice $g/0$. Therefore, when we refine (21) and (20) to direct decompositions with indecomposable factors—we denote these decompositions by (21′) and (20′)—we can replace the indecomposable element a_1 in (21′) by an indecomposable factor c_1 of (20′), i.e.,

$$g = c_1 \times d. \tag{23}$$

For the sake of definiteness, let $c_1 \leqq a_1^1$.

From (23), (22), and the inequality $c_1 \leqq g$, it follows that

$$0 = c_1 \cap d = c_1 \cap g \cap (a_2 \times \ldots \times a_k) = c_1 \cap (a_2 \times \ldots \times a_k),$$

i.e., the union

$$h = c_1 \cup (a_2 \times \ldots \times a_k)$$

is direct:

$$h = c_1 \times a_2 \times \ldots \times a_k.$$

Since a_1 and c_1 are directly similar elements in g, it follows from (10) of § 27.7 that $l(a_1) = l(c_1)$, so that, by (11) of § 27.7, $l(h) = l(1)$, and hence $h = 1$.

Thus,

$$1 = c_1 \times a_2 \times \ldots \times a_k.$$

However, $c_1 \leqq a_1^1 \leqq b_1$, and as b_1 is an indecomposable element, we have, by § 27.6,

$$c_1 = a_1^1 = b_1 \tag{24}$$

and therefore

$$1 = b_1 \times a_2 \times \ldots \times a_k.$$

We see, then, that the element a_1 in the direct decomposition (16) can be replaced by b_1; a_1 and b_1 are therefore directly similar elements. We note, moreover, that by (24) the component of a_1 in the factor b_1 of (17) coincides with b_1.

Let us now show that in the case under consideration the element b_1 can be replaced, in turn, by a_1. In fact, since by (24) $a_1^1 = b_1$, that is, by (15),

$$b_1 \cap (a_1 \cup \bar{b}_1) = b_1,$$

we have

$$b_1 \leqq a_1 \cup \bar{b}_1,$$

and therefore

$$a_1 \cup \bar{b}_1 = 1.$$

On the other hand, as a_1 and b_1 are directly similar elements, it follows that $l(a_1) = l(b_1)$, so that, by (18) of § 26.11 and (10) of § 27.7,

$$l(1) = l(b_1) + l(\bar{b}_1) = l(a_1) + l(\bar{b}_1),$$

and hence

$$a_1 \cap \bar{b}_1 = 0.$$

This proves the existence of the direct decomposition

$$1 = a_1 \times \bar{b}_1 = a_1 \times b_2 \times \ldots \times b_l.$$

14. Let us now assume that the components of a_1 in all the factors b_i of (17) coincide with the b_i and that at the same time the components in the factor a_1 of (16) of at least one b_i—say, b_1—coincides with a_1. In this case we know from the arguments of the preceding paragraph that

$$a_1 \cup \bar{b}_1 = 1, \qquad b_1 \cup \bar{a}_1 = 1. \tag{25}$$

From (16), (17), and (25) it follows, by (18) of § 26.11 and (10) of § 27.7 that

$$l(1) = l(a_1) + l(\bar{a}_1) = l(b_1) + l(\bar{b}_1),$$
$$l(1) \leqq l(a_1) + l(\bar{b}_1), \quad l(1) \leqq l(b_1) + l(\bar{a}_1).$$

Putting these equations and inequalities together, we obtain

$$l(a_1) = l(b_1), \qquad l(\bar{a}_1) = l(\bar{b}_1),$$

and consequently,

$$a_1 \cap \bar{b}_1 = b_1 \cap \bar{a}_1 = 0,$$

i.e.,

$$1 = a_1 \times b_2 \times \ldots \times b_l = b_1 \times a_2 \times \ldots \times a_k.$$

15. It remains to consider the case where the components of a_1 in all the factors b_i of (17) coincide with the b_i but the components of all the b_i ($i = 1, 2, \ldots, l$) in the factor a_1 of (16) are different from a_1. In that case all the direct factors of (17) are in the same situation as the element a_1 was in the first case discussed above. The factors b_i ($i = 1, 2, \ldots, l$) can therefore be replaced successively in (17) by factors of (16). As we know

from § 27.12, in this consecutive replacement there is a b_i that can be replaced by a_1. But then, as was shown in § 27.13, the component of b_i in the direct factor a_1 of (16) must coincide with a_1, contrary to hypothesis.

This contradiction shows that the last case is altogether impossible. //

§ 28. Direct Decompositions of Ω-groups

1. As we already know, the lattice of normal subgroups of an arbitrary group, the lattice of ideals of an arbitrary ring, and more generally, the lattice of ideals of an arbitrary Ω-group are both modular and complete. When we apply the definitions of § 27.2 to these lattices and choose as our a the group, ring, or Ω-group itself (i.e., the unit element of the lattice under consideration), then we arrive at the definition of the decomposition of a group into the *direct product of normal subgroups* and the decomposition of a ring or an Ω-group into the *direct sum of ideals.*

2. These lattice-theoretical definitions can be rephrased in the language of operations on groups, rings, or Ω-groups. We shall do this now for Ω-groups.

Let B_i $(i \in I)$ be Ω-subgroups of an Ω-group G. We denote by $\bar{B}_i$ the Ω-subgroup generated in G by all the B_j, where $j \in I$, $j \neq i$. The Ω-group G is the *direct sum of its Ω-subgroups B_i $(i \in I)$* if the following conditions α) and β) are satisfied:

α) The commutator group (see § 21.1) of the Ω-subgroups B_i and $\bar{B}_i$ is zero for all $i \in I$,

$$[B_i, \bar{B}_i] = O \qquad (i \in I). \tag{1}$$

We mention at once that condition α) implies the equation

$$[B_i, B_j] = O \qquad \text{for} \quad i \neq j. \tag{2}$$

Consequently, in particular, all commutators $[b_i, b_j]$, where $b_i \in B_i$, $b_j \in B_j$, are zero; that is, *arbitrary elements, taken one each from the subgroups B_i and B_j, with $i \neq j$, commute with each other.*

β) Every non-zero element $a \in G$ has a unique representation (to within the order of the terms) as a sum of a finite number of non-zero elements taken one each from some of the Ω-subgroups B_i, i.e.,

$$a = b_1 + b_2 + \ldots + b_k, \tag{3}$$

where $b_l \in B_{i_l}$ $(l = 1, 2, \ldots, k)$ and $i_l \neq i_m$ for $l \neq m$.

3. Let us prove the *equivalence of these two definitions.* Suppose that an Ω-group G is the direct union of its ideals B_i $(i \in I)$ in the sense of § 27.2. Then, by § 18.4, $\overline{B}_i$ $(i \in I)$ is also an ideal. Therefore, by § 21.3,

$$[B_i, \overline{B}_i] \subseteq [B_i, G] \subseteq B_i,$$
$$[B_i, \overline{B}_i] \subseteq [G, \overline{B}_i] \subseteq \overline{B}_i,$$

and since $B_i \cap \overline{B}_i = O$ by (5) of § 27.2, condition α) is proved.

Furthermore, equation (4) of § 27.2 (where G itself is taken for a) and the last theorem of § 18.4 show that for every element $a \in G$ there exists an expression of the form (3). This is a consequence of the permutability of elements taken from distinct subgroups B_i, which follows from α). Suppose that there are two such expressions, (3) and

$$a = b_1' + b_2' + \ldots + b_k';$$

we may, if necessary, add terms equal to zero to both these expressions; hence it is no loss of generality to assume that they have the same number of terms and that b_l' $(l = 1, 2, \ldots, k)$ is contained in the same Ω-subgroup B_{i_l} as b_l. If here, for example, $b_1 \neq b_1'$, that is, $-b_1' + b_1 \neq 0$, then

$$-b_1' + b_1 = b_2' + \ldots + b_k' - (b_2 + \ldots + b_k) \in B_{i_1} \cap \overline{B}_{i_1},$$

in contradiction to the equation (5) of § 27.2.

Now let the Ω-group G be the direct sum of its Ω-subgroups B_i $(i \in I)$ in the sense of the definition of § 28.2. We shall show that every Ω-subgroup B_i is an ideal in G.

If $x \in B_i$ and $a \in G$, then a has an expression of the form (3), and since elements from distinct B_j $(j \in I)$ are permutable with each other, we have

$$-a + x + a = -b_i + x + b_i \in B_i;$$

this proves that B_i is a normal subgroup of the additive group.

Furthermore, it follows from (1), in virtue of (2) of § 21.1, that for every n-ary operation $\omega \in \Omega$ and arbitrary elements $b_1, b_2, \ldots, b_n \in B_i$ and $\overline{b}_1, \overline{b}_2, \ldots, \overline{b}_n \in \overline{B}_i$

$$(b_1 + \overline{b}_1)(b_2 + \overline{b}_2) \ldots (b_n + \overline{b}_n)\omega = b_1 b_2 \ldots b_n \omega + \overline{b}_1 \overline{b}_2 \ldots \overline{b}_n \omega. \quad (4)$$

Now suppose given an n-ary operation $\omega \in \Omega$, elements $a_1, a_2, \ldots, a_n \in G$, an element $b \in B_i$, and a number k $(1 \leqq k \leqq n)$. By β), there exists an expression

$$a_j = b_j + \overline{b}_j, \quad b_j \in B_i, \quad \overline{b}_j \in \overline{B}_i \qquad (j = 1, 2, \ldots, n). \quad (5)$$

Using (4), (5), and the fact that elements from the subgroups B_i and $\bar{B}_i$ are permutable, we obtain

$$
\begin{aligned}
&-a_1 \dots a_n\omega + a_1 \dots a_{k-1}(b + a_k)a_{k+1} \dots a_n\omega = \\
&\quad = -(b_1 + \bar{b}_1) \dots (b_n + \bar{b}_n)\omega + \\
&\quad + (b_1 + \bar{b}_1) \dots (b_{k-1} + b_{k-1})(b + b_k + \bar{b}_k)(b_{k+1} + \bar{b}_{k+1}) \dots (b_n + \bar{b}_n)\omega = \\
&\quad = -\bar{b}_1 \dots \bar{b}_n\omega - b_1 \dots b_n\omega + b_1 \dots b_{k-1}(b + b_k)b_{k+1} \dots b_n\omega + \\
&\qquad\qquad + \bar{b}_1 \dots \bar{b}_n\omega \in B_i.
\end{aligned}
$$

This shows that all the B_i ($i \in I$) are ideals in G. From the condition β) it follows further that the ideals B_i together generate the whole Ω-group G. Finally, if the intersection $B_i \cap \bar{B}_i$ contained a non-zero element a, then a would have two distinct expressions of the form (3), in contradiction to β): in one of them, a is selected from B_i and zeros from the other subgroups; in the other, a is represented as a sum of elements taken one each from several subgroups B_j, where $j \neq i$. //

4. *For groups, the condition* α) *in the definition of* § 28.2 *can be replaced by the following condition*:

α') Arbitrary elements taken one each from any two subgroups B_i and B_j with $i \neq j$ are permutable.

For rings, the condition α) *can be replaced by the following condition*:

α'') The product of arbitrary elements taken one each from any two subrings B_i and B_j with $i \neq j$ is zero.

From what has been said in § 21.1 and the inclusion

$$[B_i, B_j] \subseteq [B_i, \bar{B}_i] \qquad (i \neq j)$$

it follows that α) implies α') and α''), respectively, in the two cases. On the other hand, in the case of groups it follows from α') that elements from B_i and $\bar{B}_i$ are permutable, so that—again by § 21.1—α) holds. In the case of rings, it follows from β) that every element of $\bar{B}_i$ has an expression of the form (3) with terms taken from several of the subgroups B_j ($j \neq i$). Therefore it follows from α''), by the distributive law, that if $x \in B_i$, $y \in \bar{B}_i$, then

$$xy = yx = 0,$$

and by § 21.1, this in turn implies α). //

5. Let A and B be directly similar normal subgroups of a group G in the sense of § 27.10, so that there exists a normal subgroup C and direct decompositions

$$G = A \times C = B \times C. \tag{6}$$

Then *A and B are isomorphic to the factor group G/C, i.e., they are isomorphic to each other.* In this isomorphism, those elements $a \in A$ and $b \in B$ are associated for which

$$a = bc \qquad (c \in C)$$

and *c belongs to the center of G* (see § 19.2); the isomorphism between A and B is therefore called *central.*

For in virtue of the second of the decompositions (6) the element c is permutable with each element of B. On the other hand, for every $c' \in C$ the first of the decompositions (6) implies the equation

$$ac' = c'a,$$

or

$$(bc)c' = c'(bc),$$

and from this, by what has been said above, we have

$$cc' = c'c.$$

The element c is therefore permutable with the elements of B as well as those of C, and hence with all the elements of G. //

Similarly, if A and B are directly similar ideals of a ring R, i.e., if there are direct-sum decompositions

$$R = A + C = B + C, \tag{7}$$

then *A and B are isomorphic to the factor ring R/C and are therefore isomorphic to each other.* This isomorphism associates those elements $a \in A$ and $b \in B$ for which

$$a = b + c \qquad (c \in C) \tag{8}$$

where *the element c belongs to the annihilator of R;* it is therefore called an *annihilator isomorphism.* The *annihilator* of a ring R is the set of those elements $a \in R$ for which

$$ax = xa = 0$$

for all $x \in R$.

Let us show that the element c in (8) has these properties. From the first of the decompositions (7) it follows that for all $b' \in B$, $c' \in C$

$$b'c' = c'b' = 0, \tag{9}$$

and in particular,

$$b'c = cb' = 0. \tag{10}$$

On the other hand, from the first of the decompositions (7) it follows that for every $c' \in C$

$$ac' = c'a = 0,$$

or, by (8),

$$(b + c)c' = c'(b + c) = 0,$$

and hence, by (9),

$$cc' = c'c = 0. \tag{11}$$

From (10) and (11) it follows, in virtue of the second of the decompositions (7), that c belongs to the annihilator of R. //

From what we have shown it follows that directly similar direct decompositions in groups and rings (see § 27.10) are *centrally isomorphic* and *annihilator isomorphic*, respectively. The converse, of course, does not hold.

6. The Schmidt-Ore Theorem (see § 27.12) now yields the following theorems.

If a group G has principal series, then any two of its direct decompositions with indecomposable factors are centrally isomorphic.

If a ring R has principal series, then any two of its direct decompositions with indecomposable summands are annihilator isomorphic.

❖ If all the homomorphic images of a group (ring) G lying in its center (in its annihilator) satisfy the minimal condition for subgroups (subrings), then any two direct decompositions of G, where the number of direct factors (summands) can also be infinite, have centrally (annihilator) isomorphic refinements. The conditions of this theorem are satisfied, in particular, when the center (annihilator) of G or the factor group of G with respect to its derived group (the factor ring with respect to its square) satisfies the minimal condition [A. G. Kurosh, *Izv. Akad. Nauk SSSR, ser. mat.*, vol. 10 (1946), pp. 47-72]. ❖

§ 29. Complete Direct Sums of Universal Algebras

1. Let G_i be a collection of universal algebras all with the same system of operations Ω belonging to the same primitive class A (see § 22.3); the index i ranges over a set I, finite or infinite. We shall investigate the set G whose elements are all possible systems $a = (a_i)$ of elements taken one each from every algebra G_i, i.e., $a_i \in G_i$ $(i \in I)$. The element a_i is called the *i-th component* (or the component in G_i) of a.

The set G can be made into a universal algebra with the system of operations Ω by stipulating that the operations of Ω are carried out in G *componentwise*: if an n-ary operation $\omega \in \Omega$ and n elements of G are given,

$$a^{(k)} = (a_i^{(k)}) \qquad (k = 1, 2, \ldots, n),$$

then

$$a'a'' \ldots a^{(n)}\omega = (a_i' a_i'' \ldots a_i^{(n)}\omega).$$

This definition of the operations shows that all the identical relations of A are satisfied in the algebra G.

The algebra G is called the *complete direct sum* of the algebras G_i $(i \in I)$ and is written in the form

$$G = \widetilde{\sum_{i \in I}} G_i. \tag{1}$$

We can speak, in particular, of a group being the complete direct sum (or the complete direct product) of a given collection of groups, and so forth.

2. A subalgebra A of G represented in the form (1) is called *subdirect sum* of the algebras G_i $(i \in I)$ if for every $i \in I$ the i-th components of all the elements of A fill out the whole algebra G_i. Clearly, the algebra G itself is such a subalgebra, but in general there also exist proper subalgebras of G with these properties.

If A is a subdirect sum of the algebras G_i $(i \in I)$, then for every $i \in I$ we obtain a homomorphic mapping φ_i of A onto G_i by assigning to every element

$$a = (a_i) \in A$$

its i-th component a_i. As we know from § 17.8, the homomorphism φ_i determines a congruence π_i on A, and moreover, π_i is such that the algebras G_i and A/π_i are isomorphic.

The intersection of the congruences π_i $(i \in I)$ (see § 3.3) *is the null congruence, that is, the partition of* A *into individual elements.*

For if two distinct elements are given in A, then they have distinct i-th components for at least one i, and therefore they belong to distinct classes of the congruence π_i. //

3. The converse theorem is also true:

If π_i $(i \in I)$ *is a system of congruences, with null intersection, in a universal algebra* A, *then* A *is isomorphic to a subdirect sum of the factor algebras* A/π_i $(i \in I)$.

To prove this, we let G denote the complete direct sum of the algebras A/π_i $(i \in I)$:

$$G = \widetilde{\sum_{i \in I}} A/\pi_i.$$

We map A into G by assigning to every element $a \in A$ that element of G whose i-th component (for every $i \in I$) is the congruence class π_i containing the element a. This mapping is a homomorphism, because in G the operations are performed componentwise, whereas in A/π_i $(i \in I)$ the application of the operations to the congruence classes π_i is defined by applying operations to the elements of A, taken one each in these classes.

The homomorphism so obtained is even an isomorphism, because the condition that the intersection of the congruences π_i must be null implies that for any two distinct elements of A we can find at least one $i \in I$ such that these elements lie in distinct congruence classes π_i, so that their images in G have distinct i-th components. Finally, that subalgebra of G onto which A is mapped isomorphically is a subdirect sum of the algebras A/π_i $(i \in I)$, because for all $i \in I$ every congruence class π_i, i.e., every element of A/π_i, appears as the i-th component of the images of all its elements in G. //

In the case of Ω-groups, congruences are given by ideals, as we know from § 18.5, and it is easy to see that the intersection of the congruences is determined by the intersection of the corresponding ideals. Thus, *the representations of an arbitrary Ω-group (in particular, a group or a ring) as a subdirect sum are in one-to-one correspondence with the systems of its ideals (normal subgroups) with intersection zero.*

4. A universal algebra A is called *subdirectly indecomposable* if every system of non-null congruences of A has a non-empty intersection. From the results of §§ 29.2 and 29.3 it follows that this is the case if and only if for every representation of A as a subdirect sum of suitable algebras G_i $(i \in I)$ the homomorphism φ_i that maps every element of A into its i-th component is an isomorphism between A and G_i, for at least one i.

The following theorem holds [G. Birkhoff, *Bull. Amer. Math. Soc.*, vol. 50 (1944), pp. 764-68]:

Every universal algebra A splits into the subdirect sum of pairwise indecomposable algebras belonging to the same primitive class as A.

To prove this, we examine an arbitrary pair of distinct elements $x, y \in A$ and we denote by $M(x, y)$ the set of all congruences on A that *separate* these elements, that is, those for which x and y belong to distinct classes in each of these congruences. The set $M(x, y)$ is not empty, because it contains the null congruence. Let us show that this set, partially ordered in accordance with § 2.1, has maximal elements.

By the Kuratowski-Zorn Theorem (see § 6.3), it is sufficient for this purpose to show that every chain of sets $M(x, y)$ has an upper bound. Suppose that the congruences $\pi_i \in M(x, y)$ (where i ranges over an ordered set I) form a chain, that is, $\pi_i < \pi_j$ for $i < j$. Then the binary relation π^* defined by the condition that $b\pi^* c$ if and only if there exists an $i \in I$ for which $b\pi_i c$, is obviously an equivalence relation. It is even a congruence: if elements $b_1, b_2, \ldots, b_n$ and $c_1, c_2, \ldots, c_n$ are chosen in A with $b_j \pi^* c_j$ $(j = 1, 2, \ldots, n)$, then there exist $i_1, i_2, \ldots, i_n \in I$ such that $b_j \pi_{i_j} c_j$ $(j = 1, 2, 3, \ldots, n)$. If i_0 is the largest of the indices $i_1, i_2, \ldots, i_n$, then $b_j \pi_{i_0} c_j$ $(j = 1, 2, \ldots, n)$. Let ω be an arbitrary n-ary operation given in A. Since π_{i_0} is a congruence, we have

$$(b_1 b_2 \ldots b_n \omega) \pi_{i_0} (c_1 c_2 \ldots c_n \omega),$$

and therefore

$$(b_1 b_2 \ldots b_n \omega) \pi^* (c_1 c_2 \ldots c_n \omega)$$

as well, so that π^* is, in fact, a congruence. This congruence separates the elements x and y, because they are separated by each of the congruences π_i $(i \in I)$, and it is obviously an upper bound for the given chain of congruences.

For each pair of distinct elements $x, y \in A$ let us now fix one of the maximal congruences separating them; we denote it by $\pi(x, y)$. The intersection of all these congruences is empty, because every pair of distinct elements of A is separated by at least one of these congruences. Therefore, by the theorem of § 29.3, the algebra A is a subdirect sum of all the factor algebras $A/\pi(x, y)$.

It remains to show that every factor algebra $A/\pi(x, y)$ is subdirectly indecomposable. For this purpose, we observe that there is a natural one-to-one correspondence between all the congruences of this factor algebra and all those congruences of A that contain the congruence $\pi(x, y)$. However, the intersection of all congruences of A strictly larger than $\pi(x, y)$ is itself strictly larger than this congruence, because all these congruences do not separate the elements x and y, whereas $\pi(x, y)$ does. Hence it follows that the intersection of all non-null congruences of the algebra $A/\pi(x, y)$ is itself non-null. //

5. Let us now examine the complete direct sum of Ω-groups G_i $(i \in I)$,

$$G = \sum_{i \in I}^{\sim} G_i.$$

From what has been said in § 29.1, it follows that G itself is an Ω-group of the same primitive class to which all the G_i $(i \in I)$ belong. Its zero is the

element in which every component is equal to the zero of the corresponding Ω-group G_ι.

The subset G' of G consisting of all those elements $a = (a_\iota)$ in which only a finite number of components a_ι $(\iota \in I)$ is different from zero is an Ω-subgroup of G, because the componentwise application of addition or subtraction to the elements of G', and the operation of Ω as well, do not lead outside of G'.

The Ω-group G' is called the *direct sum* of the Ω-groups G_ι $(\iota \in I)$. The justification for this is the following theorem.

The Ω-group G' is the direct sum (in the sense of the definitions of §§ 27.2 and 28.2) of ideals G_ι' $(\iota \in I)$ that are isomorphic to the corresponding Ω-groups G_ι.

To prove this, we denote by G_ι' the subset of those elements of G' in which all the components, except possibly the ι-th, are equal to zero. Since the operations are performed componentwise, this G_ι' is an Ω-subgroup of G' and even an ideal, and it is obviously isomorphic to the Ω-group G_ι.

We note, further, that every element of G' has only a finite number of non-zero components and can therefore be represented as a sum of a finite number of elements each of which has only a single non-zero component. Hence it follows that G' is the sum (see § 18.4) of all the ideals G_ι' $(\iota \in I)$. Furthermore, the sum $\overline{G}_\iota'$ of all ideals G_j', where $j \in I$ and $j \neq \iota$, consists of precisely those elements of G' whose ι-th component is equal to zero. Finally, it is clear that

$$G_\iota' \cap \overline{G}_\iota' = O,$$

so that all requirements of the definition of § 27.2 are satisfied. //

6. We therefore have three equivalent definitions for the direct sum of Ω-groups: the lattice-theoretical definition of § 27.2, the 'internal' definition of § 28.2, and the 'external' definition of § 29.5, which gives a construction that enables us to speak of the direct sum of arbitrary preassigned Ω-groups. Observe that the concept of the complete direct sum of an infinite number of direct summands is not a lattice-theoretical one.

§ 30. Distributive Lattices

1. In § 26.1, a lattice L was called distributive if the equation

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c) \tag{1}$$

is satisfied identically in it. Let us show that *the identity* (1) *is equivalent to the identity*

$$a \cup (b \cap c) = (a \cup b) \cap (a \cup c) \tag{2}$$

dual to it.

For when we apply (1) and the definition of a lattice in § 25.2, we obtain

$$a \cup (b \cap c) = [a \cup (a \cap c)] \cup (b \cap c) = a \cup [(a \cap c \cup (b \cap c)] =$$
$$= a \cup [(a \cup b) \cap c] = [(a \cup b) \cap a] \cup [(a \cup b) \cap c] = (a \cup b) \cap (a \cup c).$$

The dual arguments enable us to derive (1) from (2).

If L is a distributive lattice, then for arbitrary elements a, b, c of L it follows from

$$a \cap c = b \cap c, \qquad a \cup c = b \cup c$$

that $a = b$. In fact,

$$a = a \cup (a \cap c) = a \cup (b \cap c) = (a \cup b) \cap (a \cup c) =$$
$$= (a \cup b) \cap (b \cup c) = b \cup (a \cap c) = b \cup (b \cap c) = b.$$

❖ The property of distributive lattices just proved can also be taken as their definition. ❖

2. As we know from § 1.1, the lattice of all subsets of an arbitrary set is distributive. The same is true of arbitrary sublattices of such a lattice or, as we shall say, arbitrary *lattices of sets*. Now it turns out that, essentially, these are all the distributive lattices.

Every distributive lattice is isomorphic to a lattice of sets.

3. Distributive lattices form a primitive class of universal algebras, so that the theorem of § 29.4 is applicable to them: every distributive lattice is a subdirect sum of subdirectly indecomposable distributive lattices.

An example of a subdirectly indecomposable distributive lattice is the lattice T consisting of two elements: the null and the unit element. That this lattice is distributive, is obvious. If T were subdirectly decomposable, then it would be a sublattice of a complete direct sum of lattices T_i $(i \in I)$ and there would exist homomorphisms φ_i: $T \to T_i$ $(i \in I)$ that are not isomorphisms. But in that case, every lattice T_i $(i \in I)$ would consist of a single element, and then the direct sum of these lattices would also have only one element and could not contain T.

Every subdirectly indecomposable distributive lattice consisting of more than one element is isomorphic to T.

For let L be a distributive lattice. If it consists of two elements, then it is obviously isomorphic to T. But if L contains not fewer than three elements, then we can find an element a, different both from the null and the unit element (if L indeed has a null or a unit element). We denote by

U the sublattice of all the elements x of L satisfying the condition $x \leqq a$, and by V the sublattice of those $x \in L$ for which $x \geqq a$. By the conditions imposed on the choice of a, each of the sublattices U and V consists of more than just the single element a.

Let $U + V$ denote the complete direct sum of the lattices U and V in the sense of § 29.1: its elements are pairs (u, v), $u \in U$, $v \in V$, and the operations on them are carried out componentwise. To every $x \in L$ we assign the pair $(x \cap a, x \cup a) \in U + V$. This mapping of L into $U + V$ is one to one, because for $x, y \in L$ it follows from

$$x \cap a = y \cap a, \qquad x \cup a = y \cup a$$

that $x = y$, as was proved in § 30.1.

The mapping is even an isomorphism, because

$$((x \cup y) \cap a,\ (x \cup y) \cup a) = ((x \cap a) \cup (y \cap a),\ (x \cup a) \cup (y \cup a)) = \\ = (x \cap a, x \cup a) \cup (y \cap a, y \cup a);$$

$$((x \cap y) \cap a,\ (x \cap y) \cup a) = ((x \cap a) \cap (y \cap a),\ (x \cup a) \cap (y \cup a)) = \\ = (x \cap a, x \cup a) \cap (y \cap a, y \cup a).$$

We may therefore look upon L as a sublattice of $U + V$. Since to an element $x \in U$ there corresponds the pair (x, a) and to an element $x \in V$ the pair (a, x), we see that the homomorphism that maps each element of L into its first component and the one that maps each element into its second component map L onto the whole of U and the whole of V, respectively, so that they are not isomorphisms. Consequently, the lattice L proves to be subdirectly decomposable. //

4. We shall now prove the main theorem of § 30.2. If L is a distributive lattice, then, by what has been proved above, it is a subdirect sum of lattices T_i $(i \in I)$, where every T_i consists of two elements: the null 0_i and the unit element 1_i (one-element direct summands can, of course, be excluded from the discussion). Let M denote the set of all the unit elements 1_i $(i \in I)$ and let us assign to every element $a \in L$ the subset A of M consisting of all those unit elements 1_i that are i-th components in the expression of a as an element of the complete direct sum of the lattices T_i $(i \in I)$. It is clear that if $a \neq b$, then the sets corresponding to them are distinct. Furthermore, the fact that the operations in the complete direct sum of the lattices T_i $(i \in I)$ are performed componentwise, together with the properties of the unit element and the null, show at once that the element $a \cup b$ corresponds to the set-theoretical union $A \cup B$ of the sets A and B, and the element $a \cap b$ to the set-theoretical intersection $A \cap B$ of these sets. //

We note that *if a distributive lattice L has a null and a unit element, then in the above representation the null corresponds to the empty subset of M, and the unit element to M itself.* We also note that *if L is a finite lattice, then,* as the proof of the theorem of § 31.4 shows, it is a subdirect sum of a finite number of subdirectly indecomposable lattices and therefore *it is isomorphic to a lattice of subsets of a finite set.*

5. A distributive lattice L with null and unit element is called a *Boolean lattice* (or *Boolean algebra*) if every element $a \in L$ has a complement $\bar{a}$, where $\bar{a} \in L$ and

$$a \cap \bar{a} = 0, \qquad a \cup \bar{a} = 1.$$

Every element $a \in L$ has a unique complement; this follows from what has been proved in § 30.1.

The lattice of all subsets of a set is a Boolean lattice, because every subset A has its set-theoretical complement $\bar{A}$. A lattice of sets (see § 30.2) containing the set-theoretical complement of each of its subsets is called a *Boolean lattice of sets.*

Every Boolean lattice L is isomorphic to a Boolean lattice of sets.

This theorem has already been proved in essence in the preceding subsection. For if the element $a \in L$ corresponds to the subset $A \subseteq M$, then the complement $\bar{a}$ corresponds to the set-theoretical complement $\bar{A}$ of A in M: the expressions for the elements a and $\bar{a}$ in the complete direct sum of lattices T_i $(i \in I)$ are such that if the i-th component in one of them is 1_i, then it is 0_i in the other, and vice versa. //

6. *Every finite Boolean lattice L is isomorphic to the lattice of all subsets of a finite set.*

For the results obtained above allow us to regard L as a Boolean lattice of subsets of a *finite* set M. Let N denote the set of all those subsets of M that are minimal non-null elements of L.

Every subset A, being an element of L, coincides with the union A_0 of all the subsets it contains that occur in N. For otherwise the intersection $A \cap \bar{A}_0$ would be a non-empty element of L (because $\bar{A}_0 \in L$), that is, it would contain at least one of the subsets occurring in N, in contradiction to the definition of the set A_0.

It is now easy to see that by assigning to every $A \in L$ the set of all elements of N contained in it we obtain a one-to-one and even an isomorphic mapping of L onto the lattice of all subsets of N. //

7. An associative ring R with a unit element is called a *Boolean ring* if all its elements are idempotent, that is, if

$$a^2 = a \qquad (a \in R). \tag{3}$$

Every Boolean ring is commutative and satisfies the identity

$$2a = 0. \tag{4}$$

Indeed, for arbitrary $a, b \in R$ it follows from (3) that

$$a + b = (a + b)^2 = a^2 + ab + ba + b^2 = a + ab + ba + b,$$

and hence

$$ab + ba = 0. \tag{5}$$

Setting $b = a$ in (5) and taking account of (3), we obtain (4), or $a = -a$, and then (5) can be rewritten as

$$ab - ba = 0,$$

which proves the commutativity. //

❖ Every Boolean lattice can be converted into a Boolean ring by setting

$$a + b = (a \cap \bar{b}) \cup (\bar{a} \cap b), \qquad ab = a \cap b.$$

Conversely, every Boolean ring can be converted into a Boolean lattice by setting

$$a \cup b = a + b - ab, \qquad a \cap b = ab.$$

In this way a one-to-one correspondence is established between Boolean lattices and Boolean rings [M. H. Stone, *Trans. Amer. Math. Soc.*, vol. 40 (1936), pp. 37-111]. ❖

CHAPTER V

OPERATOR GROUPS AND OPERATOR RINGS. MODULES. LINEAR ALGEBRAS

§ 31. Operator Groups and Operator Rings

1. As the reader knows, a real vector space is an abelian group under addition whose elements are called vectors and in which a second operation is defined—multiplication of vectors by real numbers—which is connected with addition of vectors by certain axioms. Multiplication by real numbers has a meaning also in the ring of real functions of a real variable and in the ring of real square matrices of a given order. These, and many other examples, have led to the introduction of algebraic *operator* structures.

If G is a groupoid (in particular, a semigroup or a group), then we can choose an arbitrary system of endomorphisms Σ of G (see § 19.4) and consider only those subgroupoids that are mapped into themselves by all the endomorphisms of Σ. Generalizing this idea, we denote by Σ a set composed of the elements $\alpha, \beta, \ldots$ and we say that G is a *Σ-operator* groupoid and that the set Σ is an *operator domain* of this groupoid if every element $\alpha \in \Sigma$ is associated with a certain endomorphism of G. Distinct elements of Σ may be associated with one and the same endomorphism. The operator α can be described by the same symbol as the endomorphism corresponding to it, so that for arbitrary $a, b \in G$ and every $\alpha \in \Sigma$,

$$(ab)\alpha = a\alpha \cdot b\alpha. \tag{1}$$

If G is a Σ-operator group with unit element e, then it follows from (1) that

$$e\alpha = e \qquad (\alpha \in \Sigma).$$

Thus, *an operator group is a special case of a multi-operator group* (see § 18.1): all the multi-operators are unary and each of them satisfies the condition (1).

2. A subgroupoid A of a Σ-operator groupoid G is called *Σ-admissible* if it is mapped into itself under all endomorphisms corresponding to the operators of Σ, that is, if for all $\alpha \in \Sigma$,

$$A\alpha \subseteq A.$$

An arbitrary operator $\alpha \in \Sigma$, therefore, generates a certain endomorphism in every Σ-admissible subgroupoid A, and this enables us to regard the subgroupoid as an operator groupoid with the same operator domain Σ.

The intersection, provided it is not empty, of an arbitrary system of Σ-admissible subgroupoids is, of course, itself Σ-admissible. Furthermore, we can easily show, using (1) and § 9.8, that the subgroupoid generated by a certain system of Σ-admissible subgroupoids is itself Σ-admissible.

Operator groupoids G and G' with one and the same operator domain Σ are called *Σ-operator isomorphic* if there exists an isomorphic mapping φ of G onto G', a so-called Σ-operator isomorphism, such that for arbitrary $a \in G$ and $\alpha \in \Sigma$,

$$(a\alpha)\varphi = (a\varphi)\alpha. \tag{2}$$

By use of (2) we can also define *Σ-operator homomorphisms*. We can speak, in particular, of *Σ-operator endomorphisms* and *Σ-operator automorphisms* of a Σ-operator groupoid G.

From (2) it follows that *an endomorphism φ of a Σ-operator groupoid G is a Σ-operator endomorphism if and only if it is permutable (in the semigroup of endomorphisms*: see § 19.4) *with all endomorphisms corresponding to the operators of Σ.*

We note that if the set Σ is empty or if it consists only of the identity automorphisms of the groupoid G, then the study of G as a Σ-operator groupoid is tantamount to the study of G as a groupoid without operators.

3. Let Σ be a semigroup and G a Σ-operator groupoid. We say that G is a *groupoid with the operator semigroup Σ* if for every $a \in G$ and arbitrary $\alpha, \beta \in \Sigma$,

$$a(\alpha\beta) = (a\alpha)\beta. \tag{3}$$

Note that (3) necessarily holds if Σ is taken to be a sub-semigroup of the semigroup of all endomorphisms of G.

An arbitrary set Σ can be embedded in a semigroup Γ in such a way that every Σ-operator groupoid is a groupoid with the operator semigroup Γ. Here, Σ-admissible subgroupoids are also Γ-admissible, and Σ-operator-isomorphic groupoids remain also Γ-operator-isomorphic.

For Γ we can take the set of all possible *words* of the form

$$\alpha_1\alpha_2 \ldots \alpha_n, \tag{4}$$

that is, of finite ordered systems of elements of Σ, where $n \geqq 1$ and the elements $\alpha_1, \alpha_2, \ldots, \alpha_n$ are not necessarily distinct. Defining multiplication of words by the equation

$$(\alpha_1\alpha_2 \dots \alpha_n)(\beta_1\beta_2 \dots \beta_m) = \alpha_1\alpha_2 \dots \alpha_n\beta_1\beta_2 \dots \beta_m, \tag{5}$$

we convert Γ into a semigroup. By assigning to the word (4) the endomorphism of the Σ-operator groupoid Γ that is the product of the endomorphisms corresponding to $\alpha_1, \alpha_2, \dots, \alpha_n$, we make G into a Γ-operator groupoid, and by (5), the condition (3) is satisfied. It is obvious that all the remaining statements of the theorem are also true. //

4. Let G be an additively written abelian group and R an associative ring. G is called an *abelian group with the operator ring R*, or an *R-module*, if G is an R-operator group, i.e., if for arbitrary $a, b \in G$ and $\alpha \in R$ the equation (1) holds, which is now written in the form

$$(a + b)\alpha = a\alpha + b\alpha, \tag{6}$$

and if, furthermore, for arbitrary $a \in G$ and $\alpha, \beta \in R$ the equations

$$a(\alpha + \beta) = a\alpha + a\beta \tag{7}$$

and

$$a(\alpha\beta) = (a\alpha)\beta \tag{8}$$

hold.

That this concept is a natural one can be seen from the fact that *every abelian group is a module with respect to an arbitrary subring of its ring of endomorphisms*, which holds true because, by § 19.8, the conditions (6)-(8) are satisfied in this case.

This example suggests the following additional requirement: If the ring R has a unit element ε, then we shall discuss only R-modules for which ε is the *identity operator*; in other words, to the operator ε there corresponds the identity automorphism of G,

$$a\varepsilon = a \qquad (a \in G). \tag{9}$$

R-modules with the property (9) are called *unitary*.

❖ If R is an associative ring with unit element, then the investigation of arbitrary R-modules reduces completely to that of the unitary R-modules and of R-modules in which the operators of R act *trivially*, i.e., for every a of the module and α of R,

$$a\alpha = 0.$$

Every set Σ can be embedded in an associative ring R with unit element in such a way that every Σ-operator abelian group is a unitary R-module, Σ-admissible subgroups remain R-admissible (i.e., are *submodules* of the R-module), and Σ-operator-isomorphic abelian groups remain isomorphic R-modules. ❖

5. Let us consider some examples.

For an arbitrary group G, the group of its inner automorphisms (see § 19.2) is a group of operators, because, as we know, the conditions (1) and (3) are satisfied. The normal subgroups of G constitute the admissible subgroups with respect to this operator domain.

For a group G, the group of all its automorphisms is also a group of operators; and the semigroup of all endomorphisms (see § 19.4) is a semigroup of operators. Subgroups admissible with respect to these operator domains are called *characteristic* and *fully invariant*, respectively.

A real n-dimensional vector space is a unitary module over the field of real numbers. The linear subspaces are submodules of this module and the linear transformations are operator endomorphisms (see § 31.2).

6. If a is an element in an arbitrary ring R, then the mapping $x \to xa$ ($x \in R$) is an endomorphism of the additive group of R, as is shown by the distributive law; this endomorphism is called a *right multiplication*. The additive group of R can therefore be regarded as an operator group with R itself as operator domain; the admissible subgroups then constitute the right ideals of R (see § 13.10). *Left multiplications* are defined similarly: R again emerges as an operator domain for its additive group, and in this case the admissible subgroups are the left ideals. Finally, by combining these two operator domains we obtain an operator domain for the additive group of a ring R for which only the (two-sided) ideals are admissible subgroups.

If R is an associative ring, then, regarded as a set of right multiplications, it is even a ring of operators for its additive group in the sense of § 31.4, because the conditions (7), (8), and—in the case of a ring with unit element—(9) are necessarily satisfied. But if R is regarded as a set of left multiplications, then (8) is not satisfied. In this case, the equation

$$(bc)a = b(ca),$$

which holds in R, shows that the ring anti-isomorphic to R can be regarded as an operator ring for the additive group of R: two rings R and R' are called *anti-isomorphic* if there exists an isomorphism φ between the additive groups of these rings such that for all $a, b \in R$,

$$(ab)\varphi = b\varphi \cdot a\varphi.$$

7. In defining the concept of operators for rings we can, of course, make use of the endomorphisms of the ring, in analogy with the case of groups. But the examples of multiplication by a number in the ring of functions and in the ring of matrices suggest another way: namely, the consideration

of those endomorphisms φ of the additive group of R that are permutable (in the sense of multiplication of endomorphisms) with the right and left multiplications (see § 31.6), i.e., that satisfy the condition (for all $a, b \in R$)

$$(ab)\varphi = (a\varphi)b = a(b\varphi). \tag{10}$$

This is the path we shall now follow. Thus, if Σ is a set with the elements $\alpha, \beta, \ldots$, then R is called a *Σ-operator* ring if its additive group is a Σ-operator group and if, furthermore, for arbitrary $a, b \in R$ and every $\alpha \in \Sigma$,

$$(ab)\alpha = (a\alpha)b = a(b\alpha). \tag{11}$$

The concepts of a *Σ-admissible subring*, a *Σ-admissible ideal*, a *Σ-operator isomorphism*, and a *Σ-operator homomorphism* are defined by analogy with the case of groups.

If a Σ-operator ring R has a unit element, then the endomorphisms corresponding to the operators of Σ are themselves (both right and left) multiplications by elements of R that are permutable with all the elements of R, so that all the ideals and all the one-sided ideals of R are Σ-admissible.

In fact, for all $a \in R$ and $\alpha \in \Sigma$,

$$a\alpha = (a \cdot 1)\alpha = a \cdot (1\alpha),$$
$$a\alpha = (1 \cdot a)\alpha = (1\alpha) \cdot a,$$

and hence

$$a \cdot (1\alpha) = (1\alpha) \cdot a \qquad (a \in R). \text{ //}$$

Without developing this further, we mention that the definition of a Σ-operator ring can easily be carried over to the case of an arbitrary multi-operator group.

8. *Those endomorphisms of the additive group of a ring R that satisfy* (10) *form a subring in the ring of all endomorphisms of this group.* For if two endomorphisms φ and ψ have the property (10), then, for example,

$$(ab)(\varphi \pm \psi) = (ab)\varphi \pm (ab)\psi = (a\varphi)b \pm (a\psi)b = \\ = (a\varphi \pm a\psi)b = [a(\varphi \pm \psi)]b,$$

$$(ab)(\varphi\psi) = [(ab)\varphi]\psi = [(a\varphi)b]\psi = [(a\varphi)\psi]b = [a(\varphi\psi)]b. \text{ //}$$

This justifies the following concept. If Σ is an associative ring, then a ring R is called an *operator ring with Σ as ring of operators* if it is a Σ-operator ring and if, in addition, Σ is a ring of operators in the sense of § 31.4 for the additive group of R.

Observe that if n is an integer, then the mapping $a \to na$, which is defined for all elements a of R, is an endomorphism of the additive group of this ring and that the conditions (7), (8), (9), and (11) are also satisfied. Thus, *every ring R can be regarded as an operator ring with the ring of integers as ring of operators.* All the subrings of R are then admissible, and all isomorphisms are operator isomorphisms.

9. Let φ and ψ be endomorphisms of the additive group of a ring R satisfying (10). If $\varphi\psi \neq \psi\varphi$, then there exists an element $a \in R$ for which

$$a\varphi\psi \neq a\psi\varphi.$$

For every $x \in R$ we then obtain, using (10):

$$(a\varphi\psi)x = (a\varphi \cdot x)\psi = a\varphi \cdot x\psi = (a \cdot x\psi)\varphi = \\ = [(ax)\psi]\varphi = (a\psi \cdot x)\varphi = (a\psi\varphi)x.$$

Hence it follows that the non-zero element

$$b = a\varphi\psi - a\psi\varphi$$

satisfies the equation

$$bx = 0$$

for all $x \in R$, in other words, it is a *left annihilator* of R. It can also be verified that b is a *right annihilator* of R, so that it is quite generally an *annihilator*, i.e. (see § 28.5),

$$bx = xb = 0$$

for all $x \in R$.

Rings having annihilators do, of course, exist: thus, every zero ring (see § 8.2) consists of annihilators only. Nevertheless, in studying a ring with Σ as ring of operators we consider it quite justifiable to take this ring Σ to be not only associative, but also commutative.

10. Σ-operator groups (Σ-operator rings) with a given Σ form a primitive class of multi-operator groups. To what has been said in Chapter III, in particular on homomorphisms of multi-operator groups, we add only a few further remarks.

The ideals of Σ-operator groups (rings), regarded as multi-operator groups, coincide with their Σ-admissible normal subgroups (Σ-admissible ideals).

For if G is a Σ-operator group and A, a normal subgroup of G, then $aa \in A$ holds for all $a \in A$, $a \in \Sigma$ if and only if for all $a \in A$, $x \in G$, and $a \in \Sigma$,

$$(xa)^{-1}(aa)(xa) = (xa)^{-1}(ax)a \in A.$$

This last relation, however, is a rewriting, for our present case, of (2) of § 18.4.

On the other hand, if A is a subgroup of the additive group of a Σ-operator ring R, then for all $a \in A$, $x \in R$, and $\alpha \in \Sigma$, the relations $a\alpha \in A$ and

$$-x\alpha + a\alpha + x\alpha = -x\alpha + (a + x)\alpha \in A$$

are equivalent; however, the latter, again, is (2) of § 18.4. Also, we know that for a ring, regarded as a multi-operator group, the concept of an ideal coincides with that of a two-sided ideal.

11. *The operator endomorphisms of a Σ-operator abelian group G form a subring in the ring of all endomorphisms of the group* (see § 19.8). This follows at once from the definition of addition and multiplication of endomorphisms, if we use the characterization of operator endomorphisms given in § 31.2. The subring so obtained is called the *ring of operator endomorphisms* of the Σ-operator abelian group G.

Thus, the ring of operator endomorphisms of an n-dimensional real vector space (see § 31.5) is the ring of linear transformations, which, as the reader knows from his course in algebra, is isomorphic to the ring of real square matrices of order n.

12. *Let G be a group (ring) with an arbitrary operator domain Σ and H an admissible normal subgroup (admissible ideal) of G. Then the factor group (factor ring) G/H can be made into an operator group (ring) with Σ as operator domain in such a way that the natural homomorphism of G onto G/H is a Σ-operator homomorphism.*

For if G is a group, then for arbitrary $a \in G$, $h \in H$, and $\alpha \in \Sigma$,

$$(ah)\alpha = a\alpha \cdot h\alpha = a\alpha \cdot h',$$

where $h' \in H$, because H is admissible. This enables us to define the effect of the operator α on the element aH of the factor group G/H by the equation

$$(aH)\alpha = a\alpha \cdot H. \tag{12}$$

The validity of condition (1) of § 31.1 is easily verified:

$$(aH \cdot bH)\alpha = (abH)\alpha = (ab)\alpha \cdot H = (a\alpha \cdot b\alpha)H = \\ = (a\alpha \cdot H)(b\alpha \cdot H) = (aH)\alpha \cdot (bH)\alpha.$$

Thus, G/H proves to be a Σ-operator group; the last statement of the theorem follows immediately from (12).

But if G is a ring, then the equation (12), which defines the effect of the operators in G/H, can be rewritten in the form

$$(a+H)a = aa + H.$$

Let us verify the validity of the condition (11) in § 31.7:

$$[(a+H)(b+H)]a = (ab+H)a = (ab)a + H = aa \cdot b + H = \\ = (aa+H)(b+H) = (a+H)a \cdot (b+H).$$

The second half of condition (11) can be verified in the same way. //

The reader can verify without difficulty that if G is a group with Σ as semigroup of operators, or an abelian group with Σ as ring of operators, or a ring with Σ as ring of operators, then the same can be asserted about G/H, where H is a Σ-admissible normal subgroup or ideal.

§ 32. Free Modules. Abelian Groups

1. All the modules to be considered below are unitary over an associative ring R with unit element; therefore the identities (6)-(9) of § 31.4 are satisfied. All these modules form a primitive class of universal algebras (see § 22.3) with respect to the operations of the abelian group and the operators of R, so that we can speak of *free R-modules* (see § 23.2). A system of free generators of a free R-module is called a *basis*.

By making use of the definition of an abelian group and of the identities (6)-(9) mentioned above, we see at once that every element of a free R-module S with a basis X can be written in the form of a finite sum

$$x_1 a_1 + x_2 a_2 + \ldots + x_n a_n, \tag{1}$$

where $x_1, x_2, \ldots, x_n$ are distinct elements of X, $a_1, a_2, \ldots, a_n$ are non-zero elements of R, and $n \geqq 0$. Let us show that *this expression is unique*, so that we then can speak of it as *the expression* for the element in question in the basis X.

To prove this, for every $x \in X$ we form the R-module R_x isomorphic to R, regarded as a right module over itself (see § 31.6), and we take the direct sum S', in accordance with § 29.5, of all these modules R_x ($x \in X$). If we denote by x the element of R_x corresponding to the unit element of R, then by § 28.2, every element of S' has a unique expression in the form (1). The statement to be proved now follows from the fact that, by § 23.3, there exists a homomorphism of S into S' carrying every element $x \in X$ into the element x of the corresponding module R_x.

At the same time, we have shown that *every free R-module is a direct sum of a certain set of R-modules isomorphic to R itself as a right R-module, and vice versa.*

2. A right ideal B of an associative ring R with unit element is called a *principal* right ideal if there exists an element $\beta \in B$ such that every element of B can be written in at least one way in the form $\beta\alpha$ $(\alpha \in R)$. In that case, we shall write $B = (\beta)$.

If all right ideals of R are principal and R does not contain divisors of zero, then every non-zero submodule of a free R-module is itself a free R-module.

For let U be a non-zero submodule of a free R-module S with basis X. We call a set of elements $Y \subseteq U$ *admissible* if the following conditions are satisfied:

1) The submodule $\{Y\}$ generated by Y is free and has Y as a basis;

2) If X_Y is the set of all elements of X that occur in the expressions for arbitrary elements of Y in terms of the basis X, then

$$\{X_Y\} \cap U = \{Y\}.$$

We leave open for the time being the problem of whether admissible sets do exist, but we mention the fact that if they exist and if a chain of admissible sets is given, then the union of this chain is also admissible. Therefore, by the Kuratowski-Zorn Theorem (§ 6.3), among the admissible sets there exist maximal ones. Let Y be one of them.

The theorem will be proved, in virtue of 1), if we can show that $\{Y\} = U$. Let

$$\{Y\} \subset U; \tag{2}$$

this holds even when there are no admissible sets, provided we replace the submodule $\{Y\}$ by the zero submodule. We shall show that (2) leads to a contradiction.

In virtue of 2), the expression for an arbitrary element of U not lying in $\{Y\}$ must contain elements of X not lying in X_Y. Suppose that the minimal possible number of such elements is n $(n \geqq 1)$ and that in $U \setminus \{Y\}$ there exist elements in the expressions for which there occur the elements $x_1, x_2, \ldots, x_n$ of $X \setminus X_Y$ and no others. The coefficients of x_n in the expressions for all these elements, after adjunction of the zero element, form a right ideal of R and hence, by assumption, a principal right ideal (β).

Of the elements of $U \setminus \{Y\}$ under consideration we denote by z the one in the expression for which the coefficient of x_n is precisely β, and we shall show that, *in contradiction to the maximality of Y, we obtain another admissible set by adjoining z to Y.*

For if the condition 1) were not satisfied, then the elements of the submodule $\{Y, z\}$ would have non-unique expressions in terms of the elements of the set $Y \cup z$, and we could then find a non-zero element $a \in R$ such that

$$za \in \{Y\}.$$

However, this is impossible, because in the expression for z the elements $x_1, x_2, \ldots, x_n \in X \setminus X_Y$ occur with non-zero coefficients, and then, since R has no divisors of zero, they also have non-zero coefficients in the expression for za.

Furthermore, to verify the condition 2) we choose an arbitrary element $t \in \{X_{Y \cup z}\} \cap U$. In the expression for t in the basis X there occur certain elements of X_Y, and also the elements $x_1, x_2, \ldots, x_n$; here, the coefficient of x_n is contained in the right ideal (β), that is, it is of the form $\beta\gamma$ $(\gamma \in R)$. Hence it follows that the element x_n does not occur in the expression for the difference $t - z\gamma$; but then, since the number n is minimal, the elements $x_1, x_2, \ldots, x_{n-1}$ do not occur either, i.e.,

$$t - z\gamma \in \{X_Y\} \cap U = \{Y\},$$

and hence

$$t \in \{Y, z\}.$$

The proof of the theorem is now complete. //

3. Every abelian group is a module over the ring of integers, and this is a principal ideal ring (see § 15.3). By applying the results obtained above to this case and bearing in mind that the additive group of the ring of integers is an infinite cyclic group, we find:

Every free abelian group (see § 23.10) *is a direct sum of infinite cyclic groups.* The number of these cyclic direct summands (or the cardinal number of this set), which, by § 23.12, does not depend on the choice of the direct decomposition, is called the *rank* of the free abelian group.

Every non-zero subgroup of a free abelian group is itself free.

In a free abelian group S of finite rank n the rank of every subgroup is finite and does not exceed n.

For let $x_1, x_2, \ldots, x_n$ be a basis of S. We choose any k elements $y_1, y_2, y_3, \ldots, y_k$ of S, where $k > n$. Then

$$y_i = c_{i_1} x_1 + c_{i_2} x_2 + \ldots + c_{i_n} x_n \qquad (i = 1, 2, \ldots, k).$$

The system of k integral n-dimensional vectors

$$(c_{i_1}, c_{i_2}, \ldots, c_{i_n}) \qquad (i = 1, 2, \ldots, k),$$

is linearly dependent, because $k > n$, so that we can select integers l_i $(i = 1, 2, \ldots, k)$, not all zero, such that

$$l_1 y_1 + l_2 y_2 + \ldots + l_k y_k = 0.$$

Therefore the elements $y_1, y_2, \ldots, y_k$ cannot be part of the basis of any subgroup of S. //

4. Let us now prove the following main theorem on finitely generated abelian groups:

Every abelian group with a finite number of generators is a direct sum of a finite number of cyclic groups.

The proof of this theorem to be given below is due to Rado [*J. London Math. Soc.*, vol. 26 (1951), pp. 74-75].

Let G be an abelian group having a system of n generators. We examine all such systems, including systems in which some elements are equal and some are zero. Let

$$a_1, a_2, \ldots, a_n \tag{3}$$

be one of the systems of generators, where we assume, denoting the order of the element a by $o(a)$, that

$$o(a_1) \leqq o(a_2) \leqq \ldots \leqq o(a_n). \tag{4}$$

Thus, if there are zeros in (3), they occur first, and if there are elements of infinite order, they occur last.

The system (3) can be chosen in such a way that the set of numbers (4) is minimal in the lexicographic sense; that is, if $b_1, b_2, \ldots, b_n$ is another system of generators and

$$o(b_1) \leqq o(b_2) \leqq \ldots \leqq o(b_n),$$

then there is no i $(1 \leqq i \leqq n)$ for which

$$o(b_1) = o(a_1), \ldots, o(b_{i-1}) = o(a_{i-1}), \quad o(b_i) < o(a_i).$$

Let us show that *if the system of generators* (3) *is chosen in this way, then G is the direct sum of the cyclic subgroups* $\{a_1\}, \{a_2\}, \ldots, \{a_n\}$. If this were not the case, then the expression for an element of G as a sum of elements of these cyclic subgroups would not be unique, so that we could obtain an equation

$$k_1a_1 + k_2a_2 + \ldots + k_na_n = 0 \tag{5}$$

in which not all terms are equal to zero. Let

$$k_1a_1 = \ldots = k_{j-1}a_{j-1} = 0, \tag{6}$$

but $k_ja_j \neq 0$ $(j \geqq 1)$; obviously, we may assume that

$$0 < k_j < o(a_j). \tag{7}$$

We denote the greatest common divisor of the numbers $k_j, k_{j+1}, \ldots, k_n$ by k, so that

$$k_i = km_i \qquad (i = j, j+1, \ldots, n) \tag{8}$$

and the numbers $m_j, m_{j+1}, \ldots, m_n$ are relatively prime. Let us prove the following lemma.

5. *If an abelian group $A = \{a_j, a_{j+1}, \ldots, a_n\}$ and a system of relatively prime numbers $m_j, m_{j+1}, \ldots, m_n$ are given, then a new system of generators can be chosen in A,*

$$A = \{b_j, b_{j+1}, \ldots, b_n\}, \tag{9}$$

such that

$$b_j = m_j a_j + m_{j+1} a_{j+1} + \ldots + m_n a_n. \tag{10}$$

We shall prove this by induction on the sum of the absolute values

$$m = |m_j| + |m_{j+1}| + \ldots + |m_n|.$$

For $m = 1$ the statement is trivial. If $m > 1$, then in the set of relatively prime numbers $m_j, m_{j+1}, \ldots, m_n$, at least two must be different from zero. Suppose, say, that

$$|m_j| \geqq |m_{j+1}| > 0.$$

Then either $|m_j + m_{j+1}| < |m_j|$ or $|m_j - m_{j+1}| < |m_j|$, so that for one of the two signs we have

$$|m_j \pm m_{j+1}| + |m_{j+1}| + \ldots + |m_n| < m. \tag{11}$$

Since

$$A = \{a_j, a_{j+1} \mp a_j, a_{j+2}, \ldots, a_n\},$$

by taking account of (11) and applying the induction hypothesis, we obtain (9), where

$$b_j = (m_j \pm m_{j+1}) a_j + m_{j+1}(a_{j+1} \mp a_j) + m_{j+2} a_{j+2} + \ldots + m_n a_n = \\ = m_j a_j + m_{j+1} a_{j+1} + \ldots + m_n a_n.$$

6. Returning to the proof of the main theorem, we now make use of the lemma just proved and obtain for G a new system of generators

$$G = \{a_1, \ldots, a_{j-1}, b_j, b_{j+1}, \ldots, b_n\},$$

where b_j satisfies the equation (10). From (5), (6), and (8) it follows, however, that $kb_j = 0$, so that by (7)

$$o(b_j) \leqq k \leqq k_j < o(a_j),$$

in contradiction to the choice of the system of generators (3). //

7. *An infinite cyclic group cannot be decomposed into a direct sum* (see § 27.10), because by § 10.2, any two of its non-zero subgroups have a non-zero intersection.

A finite cyclic group $\{a\}$ *of order* p^n, *where* p *is a prime number, is also indecomposable,* because all its non-zero subgroups are the cyclic subgroups of the elements a, pa, p^2a, ..., $p^{n-1}a$, which are embedded in one another. Such cyclic groups are called *primary.*

A finite cyclic group $\{a\}$ *of order*

$$m = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k},$$

where $p_1, p_2, \dots, p_k$ *are distinct prime numbers and* $k \geqq 2$, *splits into the direct sum of primary cyclic groups of orders* $p_1^{n_1}, p_2^{n_2}, \dots, p_k^{n_k}$.

For let

$$q_i = p_1^{n_1} \dots p_{i-1}^{n_{i-1}} p_{i+1}^{n_{i+1}} \dots p_k^{n_k} \qquad (i = 1, 2, \dots, k).$$

Then the cyclic group $\{q_i a\}$ is of order $p_i^{n_i}$ $(i = 1, 2, \dots, k)$. By § 27.2, the sum of these subgroups is direct, because the orders of all the elements of the sum of subgroups $\{q_j a\}$ $(j \neq i)$ are prime to p_i, so that the intersection of the latter sum with $\{q_i a\}$ is zero. However, since the order of the direct sum of a finite number of finite groups is equal to the product of the orders of the direct summands, the order of the direct sum of all the subgroups $\{q_i a\}$ $(i = 1, 2, \dots, k)$ is equal to m, so that this direct sum coincides with the whole group $\{a\}$. //

By § 27.4, we derive from this the following stronger form of the main theorem:

Every abelian group with a finite number of generators is a direct sum of a finite number of indecomposable cyclic groups, some of which are infinite and the others, finite and primary.

8. In § 28.5, we introduced the concept of an isomorphism of direct decompositions (for abelian groups, we can, of course, dispense with the word 'central').

Any two decompositions of an abelian group G *with a finite number of generators into the direct sum of indecomposable cyclic groups are isomorphic.*

For let these decompositions be

$$\begin{aligned} G &= \{a_1\} + \dots + \{a_k\} + \{b_1\} + \dots + \{b_s\}, \\ G &= \{a_1'\} + \dots + \{a_l'\} + \{b_1'\} + \dots + \{b_t'\}, \end{aligned} \tag{12}$$

where the elements a are of finite order and the elements b are of infinite order. It is easy to verify that in an abelian group the set of elements of

finite order is a subgroup, the so-called *periodic part* of the group. If F is the periodic part of our group G, then

$$F = \{a_1\} + \ldots + \{a_k\} = \{a_1'\} + \ldots + \{a_l'\}, \tag{13}$$

because every element of G for which the expression in one of the decompositions (12) contains at least one of the elements b with a non-zero coefficient is necessarily of infinite order.

The group F is finite and therefore has principal series, and then the isomorphism of the direct decompositions (13) follows, by § 28.6, from the Schmidt-Ore Theorem. Furthermore, by § 28.5, the free abelian groups $\{b_1\} + \ldots + \{b_s\}$ and $\{b_1'\} + \ldots + \{b_t'\}$ of rank s and t, respectively, are isomorphic to the factor group G/F and thus to each other, and then $s = t$ (see § 32.3). //

❖ Any two decompositions of an arbitrary abelian group into the direct sum of indecomposable cyclic groups are isomorphic, provided the group has such decompositions (possibly with an infinite set of direct summands).

Every subgroup of a direct sum of cyclic groups itself splits into a direct sum of cyclic groups [L. Ja. Kulikov, *Mat. Sb.*, vol. 16 (1945), pp. 129-62]. ❖

9. An abelian group is called *primary* with respect to the prime number p if the order of each of its elements is a power of p.

An element a of a primary abelian group G (with respect to p) has finite *height* n ($n \geqq 0$) if the equation

$$p^k x = a$$

is solvable in G if and only if $k \leqq n$.

❖ Every periodic (see § 9.4) abelian group splits into the direct sum of primary groups with respect to distinct prime numbers.

A primary abelian group G splits into the direct sum of cyclic groups if and only if it is the union of an ascending sequence

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots$$

of subgroups in each of which the non-zero elements are of finite and bounded height in G [L. Ja. Kulikov, *loc. cit.*].

Every primary abelian group in which the orders of the elements are bounded splits into a direct sum of cyclic groups [Prüfer, *Math. Z.*, vol. 17 (1923), pp. 35-61].

Every countable primary abelian group whose non-zero elements are of finite height splits into a direct sum of cyclic groups [Prüfer, *loc. cit.*]. ❖

§ 33. Vector Spaces Over Fields

1. We shall investigate R-modules in the case where the associative ring R is a field. Every unitary module over an associative field K is called a *vector space* (or *linear space*) over this field. It should be remembered that the multiplication of elements of a vector space by elements of K is subject to the conditions (6)-(9) of § 31.4.

Submodules (that is, K-admissible subgroups of a vector space) are called *linear subspaces*. An *isomorphism* of a vector space over K is always to be understood, in accordance with § 31.2, as a K-operator isomorphism.

2. *Every subspace A of a vector space V over a field K has a complement, i.e., a subspace B of V for which*

$$A \cap B = O, \qquad \{A, B\} = V.$$

In other words, the space V is the direct sum of the subspaces A and B,

$$V = A + B.$$

For let us examine all those subspaces of V whose intersection with A is the zero element. The set M of all such subspaces is not empty, because it contains the null subspace. Furthermore, this set is partially ordered by set-theoretical inclusion and it satisfies the conditions of the Kuratowski-Zorn Theorem (see § 6.3).

For if L is an arbitrary chain in M, then the union C of the subspaces C_i that form this chain (i ranges over an index set) is itself a subspace of V: if $x, y \in C$, then there exist i and j such that $x \in C_i$, $y \in C_j$; if we now assume that, say, $C_i \subseteq C_j$, we have $x, y \in C_j$ and hence $x + y \in C_j$, so that $x + y \in C$; furthermore, it follows from $x \in C_i$ that $xa \in C_i$ for all $a \in K$, so that $xa \in C$. From $C_i \cap A = O$ for all i it follows, obviously, that $C \cap A = O$, and therefore the subspace C is an upper bound for the chain L in M.

Thus, M contains maximal elements. Let B be one of these maximal elements. Then $A \cap B = O$, as follows from the definition of M. Let us show that $\{A, B\} = V$.

If $x \in V$, $x \notin B$, then the subspace $\{B, x\}$ consists of all the elements of the form $b + xa$, where $b \in B$, $a \in K$. This subspace is strictly larger than B, so that by virtue of the fact that B is maximal in M, the intersection $A \cap \{B, x\}$ contains a non-zero element $a = b + xa$. Now, $a \neq 0$: otherwise we would have $a \in B$, in contradiction to the equation $A \cap B = O$. Hence we have $x = aa^{-1} - ba^{-1} \in \{A, B\}$. //

In general, a complement in V of a subspace A can be chosen in many ways. All these complements, however, are isomorphic, because by the

Isomorphism Theorem of § 20.2, or by § 28.5, they are all isomorphic to the factor space V/A.

3. The concept of linear dependence, which plays such an important role in the familiar theory of finite-dimensional vector spaces over a commutative field, can be extended to the general case that is discussed here: a finite system of elements

$$a_1, a_2, \ldots, a_k \tag{1}$$

of a vector space V over a (not necessarily commutative) field K is called *linearly dependent* if K contains elements $\alpha_1, \alpha_2, \ldots, \alpha_k$, not all equal to zero, such that

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0,$$

and *linearly independent* otherwise. As usual, linear dependence of the system (1) is equivalent to the fact that at least one of the elements in (1) can be expressed linearly in terms of the other elements of the system (with coefficients from K).

All the basic properties of linear dependence, as they are developed in an elementary course on linear algebra, carry over without difficulty to the case of arbitrary vector spaces over arbitrary associative fields. We therefore do not give proofs of these properties and confine ourselves to stating the most important of them:

If

$$a_1, a_2, \ldots, a_k$$

and

$$b_1, b_2, \ldots, b_l,$$

are two finite systems of elements in a vector space V over a field K, if the first system is linearly independent, and if each of its elements can be expressed linearly in terms of the second system, then $k \leqq l$.

Hence it follows that *if in a vector space V over a field K two finite linearly independent systems of elements are given, and if every element of each of these systems can be expressed linearly in terms of the other system, then the two systems consist of exactly the same number of elements.*

4. An arbitrary, not necessarily finite, system of elements X of a vector space V over a field K is called *linearly dependent* if K contains at least one finite linearly dependent subsystem of elements, and *linearly independent* if all finite subsystems of X are linearly independent. Clearly, every subsystem of a linearly independent system is itself linearly independent.

The space V has linearly independent subsets: for example, the subsets consisting of a single non-zero element. The set M of all linearly independent subsets of V can be partially ordered by inclusion and satisfies the condition of the Kuratowski-Zorn Theorem (see § 5.3). For if L is a chain in M, then the union X of the linearly independent subsets X_i that form this chain is itself linearly independent: if

$$x_1, x_2, \ldots, x_k \tag{2}$$

is a finite system of elements of X, then there exists in L a linearly independent subset X_j containing the whole system (2), so that the system (2) must be linearly independent. The subset X is therefore an upper bound of the chain L in M.

This shows that *a vector space V over a field K has maximal linearly independent subsets, and every linearly independent subset of this space is contained in a maximal one.*

Every maximal linearly independent subset of a vector space V is called a *basis* of the space.

5. *A subset X consisting of elements x_i (i ranges over an index set) is a basis of a vector space V over a field K if and only if every non-zero element a of V has a unique expression*

$$a = x_{i_1}\alpha^1 + x_{i_2}\alpha^2 + \ldots + x_{i_k}\alpha^k, \tag{3}$$

where $x_{i_1}, x_{i_2}, \ldots, x_{i_k} \in X$, $\alpha^1, \alpha^2, \ldots, \alpha^k$ are non-zero elements of K, and $k \geqq 1$.

For suppose that X is a basis of V, but that a non-zero element a of V does not have an expression of the form (3); from this it follows, in particular, that $a \notin X$. Then, by adjoining a to an arbitrary finite subsystem of X we obtain a linearly independent system, so that by adjoining a to the whole basis X we again obtain a linearly independent set, in contradiction to the definition of a basis. But if an element a has two distinct expressions in terms of the basis X, namely (3) and

$$a = x_{j_1}\beta^1 + x_{j_2}\beta^2 + \ldots + x_{j_l}\beta^l,$$

then the union of the finite subsystems $x_{i_1}, \ldots, x_{i_k}$ and $x_{j_1}, \ldots, x_{j_l}$ of X is linearly dependent; but this again contradicts the definition of a basis. The verification of the converse statement of the theorem is very simple and is left to the reader. //

Instead of the expression (3) we can, of course, also use, for all the elements a of the vector space V, the uniquely determined expression

$$a = \sum_{x_i \in X} x_i a^i \qquad (a^i \in K), \tag{3'}$$

in which only a finite number of the coefficients a^i is different from zero. According to § 31.4, the addition of elements written in the form (3') reduces to the addition of a coefficient for equal x_i, and the multiplication of a by an element β of K comes to the multiplication on the right by β of all the coefficients a^i in (3).

By § 34.2, we can say that *a vector space V over a field K with a basis X is the direct sum of the vector spaces xK $(x \in X)$.*

6. *All the bases of a vector space V over a field K have one and the same cardinal number.*

For let X and Y be bases of V. If at least one of them, say X, is finite and consists of m elements, then it follows immediately from § 33.3 that Y is also finite and also consists of m elements.

Suppose, then, that both bases X and Y are infinite and have the cardinal numbers m and n, respectively, where $m < n$, say. By § 33.5, every element x of X has an expression of the form (3) in terms of a finite system Y_x of elements from Y. The union Y' of all such finite subsystems Y_x, where x ranges over the whole of X, is a proper subset of Y, because, as a union of a set, with the cardinal number m, of finite sets, it also has the cardinal number m. Let y be an arbitrary element of $Y \setminus Y'$. By § 33.5, it has a linear expression in terms of a finite system of elements $x_1, x_2, \ldots, x_k$ of X and hence also in terms of the finite subsystem $\bigcup_{i=1}^{k} Y_{x_i}$ of Y'; however, this contradicts the linear independence of Y. //

The cardinal number of an arbitrary basis of a vector space V is called the *dimension* of the space. In particular, when this cardinal number is finite, we speak of a *finite-dimensional space*. Thus, the vector space xK is one-dimensional.

7. *For every cardinal number m, finite or infinite, there exists a vector space over a field K having the dimension m.*

Let X be an arbitrary set of cardinal number m. Then the direct sum of the one-dimensional vector spaces xK for all $x \in X$ (its existence follows from § 29.5) is the required vector space. Here xK is the collection of elements of the form xa $(a \in K)$ with the natural definitions of the operations of addition and multiplication by an element of K. //

Vector spaces V and W over the field K are isomorphic if and only if they have the same dimension.

For if there exists an isomorphic mapping φ of V onto W and if X is a basis of V, then $X\varphi$ is a basis of W. This proves one part of the theorem.

On the other hand, let X and Y, respectively, be bases of the spaces V and W having one and the same cardinal number. We take a fixed one-to-one mapping of X onto Y,

$$x\varphi \in Y \qquad \text{for all} \qquad x \in X,$$

and then we extend it in the following way to the whole space V: if an element a of V has the expression (3′) in terms of the basis X, then we set

$$a\varphi = \sum_{x_i \in X} (x_i\varphi)a^i.$$

We obviously obtain a one-to-one mapping of V onto the whole space W; that it is an isomorphism follows from what has been said in § 33.5. //

Thus, we see that the theorems of the present subsection give an exhaustive description of all vector spaces over an arbitrary field K.

§ 34. Rings of Linear Transformations

1. Operator endomorphisms of a vector space V over a field K are called *linear transformations* of the space. By § 31.11, linear transformations form a subring in the ring of all endomorphisms of the abelian group V. We shall denote this subring by $R(V, K)$.

We wish to show that when this ring is given, the field K as well as the space V are determined by it. To give a more accurate formulation of this statement, we introduce the following concept, which generalizes the concept of an isomorphism of two vector spaces over the same field K.

Let V and W be two vector spaces over the fields K and L, respectively. We shall say that a *semi-linear correspondence* σ is established between these spaces if there exists an isomorphism σ of K onto L and an isomorphism, which we also denote by σ, of the group V onto W, where for all $a \in V$, $\alpha \in K$

$$(a\alpha)\sigma = a\sigma \cdot \alpha\sigma.$$

Thus, if $K = L$ and σ is the identity automorphism of K, then a semi-linear correspondence becomes an isomorphism of the vector spaces V and W over K.

2. *If there is a semi-linear correspondence σ between the vector spaces V over K and W over L, then the rings of linear transformations $R(V, K)$ and $R(W, L)$ of these spaces are isomorphic.*

For let $\varphi \in R(V, K)$. Then $\sigma^{-1}\varphi\sigma$ is a mapping of W into itself, and in fact a linear transformation, because when we bear in mind that φ is a linear transformation, we obtain for arbitrary $b, b' \in W$, $\beta \in L$:

$$(b + b')(\sigma^{-1}\varphi\sigma) = b(\sigma^{-1}\varphi\sigma) + b'(\sigma^{-1}\varphi\sigma),$$

$$(b\beta)(\sigma^{-1}\varphi\sigma) = (b\sigma^{-1} \cdot \beta\sigma^{-1})\varphi\sigma = (b\sigma^{-1}\varphi \cdot \beta\sigma^{-1})\sigma = b(\sigma^{-1}\varphi\sigma) \cdot \beta.$$

The correspondence $\varphi \to \sigma^{-1}\varphi\sigma$ is a homomorphic mapping of $R(V, K)$ into $R(W, L)$, because for every $b \in W$ and arbitrary $\varphi, \psi \in R(V, K)$,

$$b[\sigma^{-1}(\varphi + \psi)\sigma] = b[\sigma^{-1}\varphi\sigma + \sigma^{-1}\psi\sigma],$$

$$b[\sigma^{-1}(\varphi\psi)\sigma] = b(\sigma^{-1}\varphi\sigma)(\sigma^{-1}\psi\sigma).$$

This is a mapping onto the whole ring $R(W, L)$, because if $\varphi' \in R(W, L)$, then

$$\varphi' = \sigma^{-1}(\sigma\varphi'\sigma^{-1})\sigma, \qquad \sigma\varphi'\sigma^{-1} \in R(V, K).$$

Finally, the correspondence is an isomorphism, because if $\sigma^{-1}\varphi\sigma$ is the null transformation, i.e., if for all $b \in W$,

$$b(\sigma^{-1}\varphi\sigma) = 0,$$

then for all $a \in V$

$$a\varphi = a\sigma(\sigma^{-1}\varphi\sigma)\sigma^{-1} = 0\sigma^{-1} = 0,$$

so that φ is also the null transformation. //

We shall say that the above isomorphism between $R(V, K)$ and $R(W, L)$ is *induced* by the semi-linear correspondence σ.

3. The converse statement also holds (see, for example, Reinhold Baer, *Linear Algebra and Projective Geometry*):

❖ Every isomorphism between the rings of linear transformations $R(V, K)$ and $R(W, L)$ is induced by some semi-linear correspondence between the spaces V and W. ❖

We shall only prove the following weaker statement.

When the ring of linear transformations $R(V, K)$ is given to within isomorphism, then it determines, also to within isomorphism, the field K as well as the space V (i.e., the dimension of this space).

4. In V we fix a basis X with the elements x_i $(i \in I)$. As usual, a linear transformation φ is completely determined, in virtue of (2) of § 31.2, when the images $x_i\varphi$ of all the elements of X are given. To every transformation φ we assign an infinite matrix whose rows and columns are numbered by all the indices $i \in I$: in the i-th column we place the coefficients of the

expression of $x_i\varphi$ in the basis X; only a finite number of these coefficients, of course, is different from zero.

Conversely, every matrix of this form, containing only a finite number of non-zero elements in each column, corresponds to some linear transformation. *The set K_I of all such matrices is a ring under the usual operations on matrices, and this ring is anti-isomorphic to $R(V, K)$* (see § 31.6). We shall only verify the following two statements:

If $A, B \in K_I$, then $AB \in K_I$. For if $i \in I$ and the i-th column of B has non-zero elements only in the rows numbered $j_1, j_2, \ldots, j_n$, then we examine the columns of A with the same numbers. They contain only a finite number of rows, so that the i-th column of the product AB only has a finite number of non-zero elements.

If $\varphi, \psi \in R(V, K)$ and if they correspond to the matrices $(\beta_{ji}), (\gamma_{kj}) \in K_I$, then the product $\varphi\psi$ corresponds to the matrix $(\gamma_{kj}) \cdot (\beta_{ji})$.

Indeed, for all $i \in I$ it follows from

$$x_i\varphi = \sum_{j \in I} x_j \beta_{ji},$$

$$x_j\psi = \sum_{k \in I} x_k \gamma_{kj}$$

that

$$x_i(\varphi\psi) = \sum_{j \in I} (x_j\psi)\beta_{ji} = \sum_{k \in I} x_k \left(\sum_{j \in I} \gamma_{kj}\beta_{ji} \right). \text{ //}$$

5. We have to show, therefore, that when the matrix ring K_I is given as an abstract ring, it determines the field K (to within isomorphism) as well as the cardinal number of the index set I.

We fix a finite system of indices

$$j_1, j_2, \ldots, j_n \in I \qquad (n \geqq 1) \tag{1}$$

and a system of non-zero elements $\lambda_t^s \in K$ $(s, t = 1, 2, \ldots, n)$ subject to the following conditions:

$$\lambda_s^s = 1; \qquad \lambda_s^t = (\lambda_t^s)^{-1}; \qquad \lambda_t^s \lambda_u^t = \lambda_u^s. \tag{2}$$

Let

$$S = S(j_1, j_2, \ldots, j_n; \lambda_t^s, s, t = 1, 2, \ldots, n) \tag{3}$$

denote the set of matrices of K_I in which only the rows whose numbers are in the system (1) can be different from zero, and where for $s, t = 1, 2, \ldots, n$, the j_s-th row is obtained by multiplying the j_t-th row on the *left* by λ_t^s.

The set S is a minimal right ideal of K_I.

The statement that S is a right ideal in K_I is easily verified. This ideal does not consist of the zero alone: in a matrix belonging to S, one row whose number is in the system (1) can obviously be given arbitrarily.

To prove that S is a minimal ideal we introduce a notation that we shall also use later: if $\alpha \in K$, then we denote by $(\alpha)_{ji}$ the matrix of K_I in which the element α occupies the position (j, i), whereas the remaining elements are equal to zero.

We choose an arbitrary non-zero matrix A in S; suppose that the element $\alpha \neq 0$ in A occupies the position (j, i), where j belongs to (1). Multiplying A on the right by the matrix $(\alpha^{-1})_{ii}$, we obtain a matrix of S with a single non-zero column, viz., the i-th, and the unit element in the position (j, i). Finally, multiplying the matrix so obtained on the right by a matrix of K_I in which only the i-th row can be different from zero, we can obviously obtain an arbitrary matrix of S. //

All the minimal right ideals of the ring K_I are of the form (3).

For let us choose an arbitrary non-zero right ideal T in K_I, and a matrix in T with a non-zero i-th column. Multiplying this matrix on the right by $(1)_{ii}$, we obtain a matrix A of T containing only a single non-zero column, the i-th. Let the non-zero elements of this column be

$$\alpha_{j_1 i},\ \alpha_{j_2 i},\ \ldots,\ \alpha_{j_n i}.$$

We set

$$\lambda_t^s = \alpha_{j_s i}\alpha_{j_t i}^{-1} \qquad (s, t = 1, 2, \ldots, n);$$

the conditions (2) are satisfied. It is now easy to see that the right ideal generated by A, which lies in T, is an ideal S of the form (3). //

6. *All the minimal right ideals of K_I, considered as right modules over K_I, are isomorphic.*

For let S and S' be two ideals of the form (3). From each of the index systems (1) associated with these ideals we select one index, j and j'. Every matrix of S is completely determined by its j-th row, which can be arbitrary; the same is true of S' and j'. We therefore obtain a one-to-one correspondence between S and S' if we assign to each other the matrices for which the j-th and j'-th rows, respectively, coincide. This correspondence is in fact an isomorphism of the K_I-modules S and S'. //

Hence it follows that *the rings of operator endomorphisms of the minimal right ideals of K_I, considered as K_I-modules, are also isomorphic.* Let us show that *these rings are anti-isomorphic to K.*

Without loss of generality we can consider an ideal S of the form (3) for which $n = 1$, so that the system (1) consists of a single index j. It is

clear that the multiplication on the left of all elements of all matrices of S by an element $\alpha \in K$ is a K_I-operator endomorphism of the ideal S.

Now let φ be an arbitrary K_I-operator endomorphism of S. Since $(1)_{jj} \in S$, we can find $(1)_{jj}\varphi$. If this matrix, in which only the j-th row can be different from zero, had a non-zero element in the position (j, i), where $j \neq i$, then we would find, in contradiction to the fact that φ is a K_I-operator endomorphism, that

$$((1)_{jj} \cdot (1)_{ii})\varphi = 0\varphi = 0,$$
$$((1)_{jj}\varphi) \cdot (1)_{ii} \neq 0.$$

Thus,

$$(1)_{jj}\varphi = (\alpha)_{jj} \qquad (\alpha \in K).$$

However, every matrix A of S is equal to its product on the left by $(1)_{jj}$, and therefore

$$A\varphi = ((1)_{jj}A)\varphi = (1)_{jj}\varphi \cdot A = (\alpha)_{jj}A.$$

But the last product signifies that all the elements of A are multiplied on the left by α.

Thus, we see that all the K_I-operator endomorphisms of S are multiplications on the left by elements of K. Furthermore, the sum of endomorphisms is associated with the sum of the corresponding elements of K and the product with the product of the elements of K, but in reverse order because multiplication by $\alpha \in K$ is performed on the left.

This shows that the field K is determined to within isomorphism by the ring $R(V, K)$ as an abstract ring. In fact, *K is anti-isomorphic to the ring of R'-operator endomorphisms of every minimal right ideal of the ring R' $(= K_I)$, anti-isomorphic to $R(V, K)$.* ///

7. We now come to the problem of the cardinal number of the index set I. Let us assume, to begin with, that the set I is finite and consists of the indices $1, 2, \ldots, n$. The set $S(j)$ of matrices of K_I in which only the j-th row can be different from zero $(j = 1, 2, \ldots, n)$ is, by § 34.5, a minimal right ideal of K_I. These ideals form a direct sum (see § 27.2), which coincides with the whole ring K_I, so that K_I, as a right K_I-module, has a composition series of length n (see § 20.6)

$$O \subset S(1) \subset S(1) + S(2) \subset \ldots \subset \sum_{j=1}^{n} S(j) = K_I.$$

In virtue of the Jordan-Hölder Theorem, the number n is therefore determined invariantly by K_I itself as a K_I-module.

We now discuss the case of an infinite set I. Let us call a system of minimal right ideals of the ring K_I *independent* if the right ideal generated

by all the ideals of this system is their direct sum. The system of ideals $S(j)$ (see above), taken for all $j \in I$, is an independent system, and even a maximal one, because every ideal S of the form (3) is contained in the direct sum of the ideals $S(j)$, where j is in (1), so that it cannot be added to the system of all ideals $S(j)$ $(j \in I)$ without destroying its independence. The cardinal number of the system of ideals $S(j)$, of course, coincides with that of the set I.

The cardinal number of every independent system of minimal right ideals of K_I does not exceed the cardinal number of I.

For if a system of ideals of the form (3) is given, then the set of index systems (1) selected in them is contained in the set of all finite subsets of the infinite set I, so that its cardinal number does not exceed the cardinal number of I. On the other hand, from the same results on composition series (see § 20.6) it follows that an independent system of ideals of the form (3) connected with the given system of indices (1) can only be finite.

Thus, the cardinal number of the set I (that is, the dimension of the vector space V) is also uniquely determined by $R(V, K)$ as an abstract ring. Specifically, *it is the maximal cardinal number of an independent system of minimal right ideals of a ring anti-isomorphic to* $R(V, K)$. //

8. If M is a set in a ring R, then we define the *right annihilator* of this set as the totality of those elements $r \in R$ for which $xr = 0$ for all $x \in M$. The right annihilator of an arbitrary set is obviously a right ideal of R. The *left annihilator* of a set is defined similarly.

❖ An associative ring R is isomorphic to the ring $R(V, K)$ of all linear transformations of a vector space V over a field K if and only if it satisfies the following conditions: 1) R has a unit element; 2) R has minimal one-sided ideals, which are all contained in every non-zero two-sided ideal; 3) If the right annihilator of a left ideal is equal to zero, then this left ideal contains all the minimal left ideals; 4) The sum of any two right (left) annihilators is itself such an annihilator [K. G. Wolfson, *Amer. J. Math.*, vol. 75 (1953), pp. 358-86]. ❖

§ 35. Simple Rings. Jacobson's Theorem

1. From what has been said in § 13.9, it follows that the complete matrix ring of given order n over an associative field K (i.e., the ring of all linear transformations of an n-dimensional vector space over K) is simple. We now wish to study a much wider class of simple associative rings.

We consider an arbitrary vector space over an associative field K. The image $V\varphi$ of this space under a linear transformation φ is obviously a linear

subspace (see § 33.1). We say that φ is a linear transformation of *finite rank* if $V\varphi$ is a finite-dimensional subspace. It is clear that the sum and difference of linear transformations of finite rank, and also the product of linear transformations at least one of which is of finite rank, are themselves transformations of finite rank; hence *the set* $R'(V, K)$ *of all such transformations is a subring, and even an ideal, of the ring of all linear transformations* $R(V, K)$.

A subring R_0 of $R(V, K)$ is called a *dense ring of linear transformations* if for every linearly independent finite system of elements $x_1, x_2, \ldots, x_n \in V$ and an arbitrary system of elements $y, y_2, \ldots, y_n \in V$ we can find a linear transformation φ in R_0 such that

$$x_i\varphi = y_i \qquad (i = 1, 2, \ldots, n).$$

2. *Every dense ring of linear transformations of finite rank is simple.*

For let R_0 be a dense ring of linear transformations of finite rank in a vector space V over a field K, and A an ideal of R_0. We shall show that *for every finite-dimensional subspace* L *we can find in* A *a projection of* V *onto* L, that is, a transtormation $a \in A$ for which $Va = L$ and for which $ya = y$ holds for every $y \in L$.

First, let L be one-dimensional: $L = xK$ $(x \neq 0)$. We choose $a \in A$ $(a \neq 0)$. The subspace Va is finite-dimensional; let $y_1, y_2, \ldots, y_n$ be one of its bases. Then there exists an $x_1 \in V$ such that $x_1 a = y_1$. Since R_0 is dense, there exists a $\varphi \in R_0$ such that

$$y_1\varphi = x_1, \qquad y_2\varphi = \ldots = y_n\varphi = 0. \tag{1}$$

Since A is an ideal, $a\varphi = a' \in A$ and, by (1), $Va' = x_1 K$, where $x_1 a' = x_1$. Further, there exist $\varphi_1, \varphi_2 \in R_0$ such that

$$x\varphi_1 = x_1, \qquad x_1\varphi_2 = x.$$

Therefore

$$V(\varphi_1 a' \varphi_2) = xK = L,$$
$$x(\varphi_1 a' \varphi_2) = x,$$

and since $\varphi_1 a' \varphi_2 \in A$, this is the required projection.

Suppose now that our statement is already proved for subspaces of dimension $n-1$ and that L is an n-dimensional subspace $(n > 1)$. We choose an $(n-1)$-dimensional subspace L' in L; by hypothesis, there exists a projection $a' \in A$ of V onto L'. Then $La' = L'$. The kernel L'' of this transformation is a one-dimensional subspace, $L''a' = O$, and

$$L = L' + L''.$$

There exists a projection $a'' \in A$ of V onto L''. Then

$$a = a' + a'' - a'a'' \in A$$

is the required projection of V onto L. In fact, for all $x \in V$ we have

$$xa = xa' + xa'' - (xa')a'' \in L;$$

if $x \in L'$, then

$$xa = x + xa'' - xa'' = x;$$

if $x \in L''$, then $xa' = 0$ and

$$xa = xa'' = x.$$

The proof of the theorem now proceeds without difficulty. For every $\varphi \in R_0$ the subspace $V\varphi$ is finite-dimensional and, by what has been proved, there exists an $a \in A$ that is a projection of V onto $V\varphi$. Therefore $a\varphi = a$ and hence $\varphi \in A$, that is, $A = R_0$. //

3. *Every dense ring of linear transformations of finite rank has minimal right ideals.*

For let R_0 again be a dense ring of linear transformations of finite rank in a vector space V over a field K and let L be an arbitrary one-dimensional subspace of V. By applying the statement proved above to the case $A = R_0$, we can assert the existence of an $a \in R_0$ that is a projection of V onto L. We denote by L' the set of those $y' \in V$ for which $y'a = 0$. Then $L \cap L' = O$ and

$$V = L + L'. \tag{2}$$

We denote by A the set of those $a' \in R_0$ for which $L'a' = O$. It is clear that $a \in A$—so that A does not consist of the zero alone—and that A is a right ideal of R_0. We shall show that it is minimal. Let a_1 and a_2 be arbitrary non-zero elements of A. If $L = xK$, then $xa_1 \neq 0$, because otherwise we would have $Va_1 = O$, that is, $a_1 = 0$. Therefore, since R_0 is dense, there exists a $\varphi \in R_0$ such that

$$(xa_1)\varphi = xa_2, \tag{3}$$

and since

$$L'a_1 = L'a_2 = O,$$

we have, by (2) and (3),

$$a_1\varphi = a_2,$$

so that the right ideal A is generated by any of its non-zero elements a_1. //

4. The converse of the above results is the following *Theorem of Jacobson* [*Trans. Amer. Math. Soc.*, vol. 57 (1945), pp. 228-45]:

Every simple ring R that is not a zero ring (see § 8.2) *and has minimal right ideals is isomorphic to a dense ring of linear transformations of finite rank of a vector space over a field.*

We choose a minimal right ideal A in our ring R. The multiplication of A on the right by an element $r \in R$ determines an endomorphism of the additive group of A. Since the sum (product) of elements of R is associated with the sum (product) of the corresponding endomorphisms, we have a homomorphism of R into the ring of endomorphisms of the additive group of A.

From the fact that R is simple, it follows that this homomorphism is either an isomorphism or the mapping into zero. However, the latter is impossible. From $AR = O$ (i.e., $ar = 0$ for all $a \in A$, $r \in R$) it would follow that $A^2 = O$, and then the non-zero two-sided ideal

$$B = R\,A + A,$$

that is, the set of elements of the form

$$\sum_{i=1}^{n} r_i a_i + a \qquad (r_i \in R;\quad a_i, a \in A;\quad i = 1, 2, \ldots, n)$$

(for all natural numbers n), would have the property $B^2 = O$. However, from the fact that R is simple it follows that $B = R$, and then $R^2 = O$, in contradiction to the fact that R is not a zero ring.

5. Consequently, we may assume that R is a ring of endomorphisms (that is, a subring of the ring of all endomorphisms) of the additive group of A. Since A is a minimal right ideal, the group A has no non-trivial subgroup that is admissible (see § 31.2) with respect to all endomorphisms of R; in other words, R is a so-called *irreducible* ring of endomorphisms.

SCHUR'S LEMMA: *If R is an irreducible ring of endomorphisms of an abelian group A, then the set K of those endomorphisms of A that are permutable (in the sense of multiplication of endomorphisms) with every endomorphism of R is a field.*

For it is clear that K is a subring of the ring of all endomorphisms of A and contains the identity automorphism of A. Now let $\alpha \in K$ and $\alpha \neq 0$. Then $A\alpha \neq O$, and since the subgroup $A\alpha$ admits multiplication on the right by all elements of R and R is an irreducible ring of endomorphisms, we have $A\alpha = A$. The kernel of the endomorphism α is an R-admissible subgroup, so that—since R is an irreducible ring—it is equal to zero. Thus,

the endomorphism α proves to be an automorphism, so that the inverse automorphism α^{-1} exists, which, like α, is permutable with all the endomorphisms of R; therefore $\alpha^{-1} \in K$. //

When we apply Schur's Lemma to the theorem to be proved, we see that the additive group of A is a vector space over a field K, and since every endomorphism of R is permutable with each endomorphism from K, R is a ring of linear transformations of this vector space.

6. *The ring R is a dense ring of linear transformations.*

First, let a and a' be elements of A with $a \neq 0$. The set aR is a right ideal of R. Suppose that $aR = O$; then the set of all $x \in A$ for which $xR = O$ does not consist of the zero alone; and since this set is a right ideal in R, it coincides with A, because A is minimal. Hence it follows that $AR = O$, but this, as was shown in § 35.4, is impossible. Therefore $aR = A$, and then there exists an $r \in R$ such that $ar = a'$.

Now let $a_1, a_2, \ldots, a_n$ be linearly independent elements over K and $a_1', a_2', \ldots, a_n'$ $(n > 1)$, arbitrary elements of A. Let us assume that the existence of elements $r_i \in R$ $(i = 1, 2, \ldots, n)$ such that

$$a_i r_i \neq 0, \qquad a_j r_i = 0 \text{ for } j = 1, \ldots, i-1, i+1, \ldots, n$$

has already been established. By what has been proved above, there exist $r_i' \in R$ $(i = 1, 2, \ldots, n)$ such that

$$a_i r_i r_i' = a_i'.$$

Therefore the element

$$r = \sum_{i=1}^{n} r_i r_i' \in R$$

has the property that

$$a_i r = a_i' \qquad (i = 1, 2, \ldots, n).$$

7. We have to show, therefore, that *for every system of elements $a_1, a_2, \ldots, a_n \in A$ linearly independent over K, there exists an element $r \in R$ such that*

$$a_1 r = \ldots = a_{n-1} r = 0, \qquad a_n r \neq 0.$$

First, let $n = 2$. We assume that it always follows from $a_1 r = 0$ that $a_2 r = 0$. For every $a \in A$ there exists an $r \in R$ such that $a = a_1 r$. If $a = a_1 r'$ also holds, then $a_1(r - r') = 0$, so that, by hypothesis, we also have $a_2(r - r') = 0$ and hence $a_2 r = a_2 r'$. A single-valued mapping α of A into itself is therefore defined if we put, for $a = a_1 r$,

$$a\alpha = a_2 r.$$

The mapping α is an endomorphism of the group A, because it follows from $a = a_1 r$, $b = a_1 r'$ that $a + b = a_1(r + r')$ and therefore that

$$(a + b)\alpha = a_2(r + r') = a_2 r + a_2 r' = a\alpha + b\alpha.$$

The endomorphism α is permutable with every endomorphism $r_0 \in R$, because it follows from $a = a_1 r$ that $ar_0 = a_1(rr_0)$, and hence

$$(ar_0)\alpha = a_2(rr_0) = (a\alpha)r_0.$$

This shows that $\alpha \in K$. Since

$$(a_1\alpha)r = (a_1 r)\alpha = a_2 r \qquad (r \in R),$$

we have

$$(a_1\alpha - a_2)r = 0 \qquad (r \in R),$$

and therefore, as we know from § 35.6, $a_1\alpha - a_2 = 0$, in contradiction to the fact that the system of elements a_1, a_2 is linearly independent over K.

8. Suppose now that n is arbitrary and that the statement of § 35.7 is already proved for $n - 1$. Therefore there exists an $r' \in R$ such that

$$a_1 r' = \ldots = a_{n-2} r' = 0, \qquad a_{n-1} r' \neq 0;$$

we denote the set of all such r' by R'. If the elements $a_{n-1} r'$ and $a_n r'$ are linearly independent over K, then, by what has been shown, there exists an $r'' \in R$ such that

$$a_{n-1} r' r'' = 0, \qquad a_n r' r'' \neq 0,$$

and we can put $r = r'r''$.

Suppose, however, that

$$a_{n-1} r' = (a_n r')\beta \qquad (\beta \in K;\ \beta \neq 0). \tag{4}$$

Again using the induction hypothesis, we can find an $r_0 \in R$ such that

$$a_1 r_0 = \ldots = a_{n-2} r_0 = 0, \qquad (a_{n-1} - a_n\beta) r_0 \neq 0. \tag{5}$$

If $a_n r_0 = 0$, $a_{n-1} r_0 \neq 0$, then we can find an $r_1 \in R$ such that $a_{n-1} r_0 r_1 = a_{n-1} r'$, and then we can put $r = r' - r_0 r_1$, because $a_n r = a_n r' \neq 0$. If $a_{n-1} r_0 = 0$, $a_n r_0 \neq 0$, then we simply have $r = r_0$. If $a_{n-1} r_0$ and $a_n r_0$ are linearly independent, then the existence of the required r has already been shown above, because $r_0 \in R'$.

It remains to discuss the case

$$a_{n-1} r_0 = (a_n r_0)\gamma \qquad (\gamma \in K;\ \gamma \neq 0); \tag{6}$$

it follows from (5) that $\beta \neq \gamma$. There exist an $r_2 \in R$ such that

$$a_{n-1} r_0 r_2 = a_{n-1} r'. \tag{7}$$

By (6), (7), and (4) we then have

$$a_n r_0 r_2 = a_{n-1} r_0 \gamma^{-1} r_2 = (a_{n-1} r_0 r_2)\gamma^{-1} = a_{n-1} r' \gamma^{-1} = a_n r' \beta \gamma^{-1},$$

and hence

$$a_n (r' \beta \gamma^{-1} - r_0 r_2) = 0. \tag{8}$$

Now we can put $r = r' - r_0 r_2$: it is clear that

$$a_1 r = \ldots = a_{n-1} r = 0;$$

if we also have $a_n r = 0$, then by (8) we would have

$$a_n r' = a_n r' \beta \gamma^{-1},$$

and since $\beta \gamma^{-1} \neq 1$, we would obtain $a_n r' = 0$, which is false.

9. *The ring R consists of linear transformations of finite rank.*

It is sufficient to show that R contains at least one non-zero linear transformation of finite rank, because then the intersection of R with the ideal $R'(A, K)$ of the ring $R(A, K)$ (see § 35.1) is a non-zero ideal of R and, since R is simple, coincides with R.

We know from § 35.4 that $A^2 \neq O$. From this, and from fact that A is a minimal right ideal, it follows that for every given non-zero $a \in A$ we have $aA = A$. There exists, in particular, an $e \in A$ $(e \neq 0)$ such that $ae = a$. Hence

$$a(e^2 - e) = 0,$$

and therefore $e^2 = e$, because the elements of A that annihilate a on the right form a right ideal. The element e is consequently *idempotent*. It is clear that $A = eR$.

If we now regard e as a linear transformation of the vector space A over the field K belonging to R, we can show that it is of finite rank. For, the subspace Ae is one-dimensional. In fact, if a' and a'' were linearly independent elements of Ae, then, since e is idempotent,

$$a'e = a', \qquad a''e = a''. \tag{9}$$

In accordance with § 35.7, there exists an $r \in R$ such that

$$a'r = 0, \qquad a''r \neq 0.$$

Hence

$$a'(er) = 0, \qquad a''(er) \neq 0,$$

so that $er \neq 0$. Since $er \in A$, the set of elements of A that annihilate a' on the right does not consist of the zero alone. This set is a right ideal and therefore coincides with A; however this contradicts the first of the equations (9). The proof of Jacobson's Theorem is now complete. //

❖ If R_i $(i = 1, 2)$ is a dense ring of linear transformations of finite rank in a vector space V_i over a field K_i and if there exists an isomorphism τ between the rings R_1 and R_2, then there exists a semi-linear correspondence σ (see § 34.1) between the spaces V_1 and V_2 such that for all $a \in V_1$, $r \in R_1$ we have

$$(ar)\sigma = a\sigma \cdot r\tau.$$

The complete rings of matrices of finite order over fields are simple associative rings satisfying the minimal condition for right ideals and, apart from zero rings, they are the only ones. ❖

§ 36. Linear Algebras. The Quaternion Algebra and the Cayley Algebra

1. When we now proceed to operator rings (see § 31.7), it is natural to discuss as a special case rings that admit as ring of operators (see § 31.8) a certain associative field. Bearing in mind what was said in § 31.9, we shall assume here that this field is commutative.

A ring R that has a field P as ring of operators is called a *linear algebra* over P. However, we usually speak simply of an *algebra* over the field, because there is no danger of confusing this concept with that of a universal algebra. If R is an associative field, then it is customary to speak of an (associative) *division algebra.*

In the non-associative case also, we define a *division algebra,* in accordance with the terminology introduced in § 12.1, as a ring with division that is an algebra over a given commutative field. The concept of an *algebra with unique division* has a corresponding meaning. But if an algebra over a commutative field is a non-commutative field, then this only signifies, as will be shown below, that a certain commutative subfield is singled out in the center of our field; however, we shall occasionally speak of an *algebra with unique division and a unit element.*

All that has been said in § 31.7 is applicable to the case of algebras. Thus, *isomorphisms* and *homomorphisms* of algebras over a field P must be understood as P-operator isomorphisms and homomorphisms. P-admissible subrings of an algebra are called *subalgebras.* An *ideal* of an algebra will always mean a P-admissible ideal. The meaning of the term *factor algebra* also requires no special explanation.

The additive group of every algebra R over a (commutative) field P is a vector space over this field, as follows immediately from the definition of an algebra. Vector spaces lend themselves to investigation much more readily than, for example, arbitrary abelian groups without operators. For this reason, the theory of algebras is in many respects simpler and considerably further developed than the parallel theory of rings without operators.

2. In § 18.9 we introduced the concept of the center of a ring; we recall that the unit element of the ring, if it exists, belongs to its center.

If a ring R has a unit element e and if a subfield P containing e is contained in the center of R, then R is an algebra over P.

For if the product of elements of R by elements of P is interpreted in the sense of the multiplication given in R, then it follows immediately from the properties of the operations in the ring and the definition of the center that all the conditions in the definition of an algebra are satisfied.

Thus, the field of complex numbers is a division algebra over the field of real numbers. The ring of polynomials $P[x]$ over a (commutative) field P is an algebra over this field, because the polynomials of degree 0 form a subfield in $P[x]$ isomorphic to P and containing the unit element of the ring. Every, not necessarily commutative, field is an algebra over its center, because the latter, as was shown in § 18.9, is a commutative field.

If an algebra R over a field P has a unit element e, then the center of the algebra has a subfield containing e and isomorphic to P.

To prove this, we let P' denote the set of elements of the form ea, where $a \in P$. In virtue of (7) and (8) of § 31.4 and (11) of § 31.7, we have

$$e(a \pm \beta) = ea \pm e\beta,$$
$$e(a\beta) = (ee)(a\beta) = [(ee)a]\beta = (ea \cdot e)\beta = ea \cdot e\beta.$$

Hence it follows that the mapping $a \to ea$ is a homomorphism of P onto P', and in fact, even an isomorphism, because P is a field and by (9) of § 31.4,

$$1 \to e \cdot 1 = e \neq 0,$$

where 1 is the unit element of P.

Finally, let us show that the subfield P' is contained in the center of R. For arbitrary $x, y \in R$, we have

$$(ea)x = (ex)a = (xe)a = x(ea),$$
$$[(ea)x]y = [(ex)a]y = xa \cdot y = (xy)a = [e(xy)]a = (ea)(xy),$$
$$(xy)(ea) = [(xy)e]a = (xy)a = x \cdot ya = x[(ye)a] = x[y(ea)].$$

This completes the proof. //

3. Let R be an algebra over a field P; in the additive vector space of this algebra let us choose a basis X, consisting of elements x_i (where i ranges over an index set). Every element a of R has a unique expression of the form (3′) of § 33.5, namely,

$$a = \sum_{x_i \in X} x_i \alpha^i \qquad (\alpha^i \in P),$$

where the addition of elements of R and their multiplication by an element β of P reduce to the addition of the corresponding coefficients and the multiplication of the coefficients by β.

If x_i, $x_j \in X$, then the product $x_i x_j$, being an element of R, also has an expression in terms of the basis X,

$$x_i x_j = \sum_{x_k \in X} x_k \varepsilon_{ij}^k, \tag{1}$$

where for given i and j only a finite number of coefficients ε_{ij}^k can be different from zero. *The system of elements ε_{ij}^k of the field P completely determines the multiplication in the algebra R.* Indeed, for arbitrary $\alpha, \beta \in P$ we have

$$(x_i \alpha)(x_j \beta) = (x_i x_j)(\alpha\beta) = \sum_{x_k \in X} x_k (\varepsilon_{ij}^k \alpha\beta).$$

If any two elements

$$a = \sum_{x_i \in X} x_i \alpha^i, \qquad b = \sum_{x_j \in X} x_j \beta^j \tag{2}$$

are given in R, then by taking into account that only a finite number of coefficients α^i and β^j can be different from zero, we obtain, on the basis of the distributive laws,

$$ab = \sum_{x_i \in X} \sum_{x_j \in X} (x_i \alpha^i)(x_j \beta^j) = \sum_{i,\, j,\, k} x_k (\varepsilon_{ij}^k \alpha^i \beta^j). \tag{3}$$

Of course, when we go over from the basis X to another basis of R, the numbers ε_{ij}^k change, so that the algebra R in this new basis is determined by a *'multiplication table'* different from (1).

4. *If a basis X consisting of the elements x_i $(i \in I)$ is chosen in a vector space V over a field P and arbitrary elements ε_{ij}^k are taken in P $(i, j, k \in I)$, then there exists an algebra over P that has V as its additive vector space and that is given in the basis X by the multiplication table* (1) *with these coefficients ε_{ij}^k.*

For if we take the formula (3) as the definition of the product of any two elements $a, b \in V$ whose expressions in the basis X are of the form (2), then

it is easy to verify that the distributive laws and the postulate (11) of § 31.7 are satisfied and that the elements of the basis X are multiplied in accordance with (1).

Clearly, the algebra constructed in this way will in general be neither associative nor commutative (that is, if the choice of the numbers ε_{ij}^{k} is not subject to additional restrictions). With the help of the distributive laws, the expression of the elements of an algebra in terms of the basis, and the properties of multiplication in a field, it is easy to verify the following statement: *If a basis X is chosen in an algebra R over a field P, then in order for this algebra to be associative, commutative, or a Lie algebra* (see § 8.3) *it is not only necessary* (*which is obvious*), *but also sufficient that for arbitrary $x_i, x_j, x_k \in X$ the following equation should hold*:

$$(x_ix_j)x_k = x_i(x_jx_k) \qquad \textit{(for an associative algebra)};$$

$$x_ix_j = x_jx_i \qquad \textit{(for a commutative algebra)};$$

$$\left.\begin{aligned} x_ix_j &= -x_jx_i, \\ (x_ix_j)x_k + (x_jx_k)x_i + (x_kx_i)x_j &= 0. \end{aligned}\right\} \qquad \textit{(for a Lie algebra)}.$$

5. Let us apply the results of § 33.7 and § 36.4 to the construction of a special type of algebra. We take an arbitrary groupoid G and an arbitrary field P and then we construct a vector space over P having the set G as a basis. By defining the product of two elements of this basis as their product in the groupoid G, we obtain an algebra, the so-called *groupoid algebra* of G over P. The reader will, of course, compare this concept with that of the integer group ring of § 10.5.

The groupoid algebra is associative if the groupoid G is a semigroup or, in particular, a group. In that case we speak of the *semigroup algebra* and the *group algebra*, respectively.

Thus, the algebra of polynomials $P[x]$ over a field P is a semigroup algebra: a basis of this algebra is the set of powers $x^0 = 1, x, x^2, \ldots, x^n, \ldots$ of the indeterminate x, which forms a semigroup under multiplication isomorphic to the additive semigroup of non-negative integers.

6. If the additive vector space of an algebra R over a field P is finite-dimensional (see § 33.6), then the algebra itself is also called *finite-dimensional*; its *dimension* is sometimes called its *rank*.

Thus, the ring of matrices P_n over a field P (see § 8.6) is, by § 36.2, an algebra over this field, because the center of the ring P_n consists of the set of scalar matrices (see § 10.3), which contains the unit matrix and is isomorphic to P. The algebra of matrices P_n then has the finite dimension n^2, because the matrices e_{ij} $(i, j = 1, 2, \ldots, n)$, containing 1 in the (i, j) position

and zeros elsewhere constitute a basis. The multiplication table of P_n in this basis is as follows:

$$e_{ij} \cdot e_{jk} = e_{ik},$$
$$e_{ij} \cdot e_{kl} = 0 \qquad \text{for } j \neq k.$$

7. As we have already mentioned in § 36.2, the field of complex numbers K is a division algebra over the field of real numbers or, as we shall say, a *real* division algebra. This algebra is of dimension 2, because the numbers 1 and i form a basis of it. In this basis, K has the following multiplication table:

$$1^2 = 1, \quad 1 \cdot i = i \cdot 1 = i, \quad i^2 = -1.$$

8. We shall now construct a four-dimensional real associative division algebra Q, the so-called *quaternion algebra*. This is an algebra with a basis $1, i, j, k$ and the following multiplication table:

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

(4)

To find a product in this table, for example, the product of i by j, we have to take the intersection of the row numbered i with the column numbered j; thus, $ij = k$.

From the multiplication table (4) it follows at once that the element 1 is the unit element of Q and that this algebra is non-commutative. Furthermore, *the algebra Q is associative*: since—apart from sign—there is complete symmetry in the way in which the elements i, j, k occur in the multiplication table (4), it is sufficient, by § 36.4, to verify the equations

$$(ii)i = i(ii),$$
$$(ii)j = i(ij), \quad (ij)i = i(ji), \quad (ji)i = j(ii),$$
$$(ij)k = i(jk);$$

this verification is left to the reader.

9. Every *quaternion* (i.e., element of Q) a has a unique expression

$$a = a + ib + jc + kd$$

with real coefficients a, b, c, d. The quaternion

$$\bar{\alpha} = a - ib - jc - kd$$

is called *conjugate* to the quaternion α. It is easy to verify that

$$\overline{\alpha\beta} = \bar{\beta}\bar{\alpha}, \qquad \overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta} \tag{5}$$

and that

$$\alpha\bar{\alpha} = \bar{\alpha}\alpha = a^2 + b^2 + c^2 + d^2. \tag{6}$$

The non-negative real number

$$n(\alpha) = \alpha\bar{\alpha} = \bar{\alpha}\alpha, \tag{7}$$

which is equal to zero for $\alpha = 0$ only, is called the *norm* of α. It is easy to verify that

$$n(\alpha\beta) = n(\alpha) \cdot n(\beta), \tag{8}$$

so that *the quaternion algebra does not contain divisors of zero*. In fact, by (7) and (5),

$$n(\alpha\beta) = (\alpha\beta)(\overline{\alpha\beta}) = \alpha\beta\bar{\beta}\bar{\alpha} = (\alpha\bar{\alpha}) \cdot n(\beta) = n(\alpha) \cdot n(\beta).$$

If $\alpha \neq 0$, and hence $n(\alpha) \neq 0$, then by (7),

$$\alpha \ ([n(\alpha)]^{-1}\bar{\alpha}) = ([n(\alpha)]^{-1}\bar{\alpha}) \cdot \alpha = 1.$$

Every non-zero quaternion α therefore has an inverse quaternion, and hence *the quaternion algebra is a field.*

10. Finally, let us construct an eight-dimensional real non-associative algebra with unique division and with the unit element, the so-called *Cayley algebra.*

We consider all possible expressions of the form $\alpha + \beta e$, where α and β are quaternions and e is a new symbol; in the particular case $\beta = 0$, instead of writing $\alpha + \beta 0$, we shall simply write α; and when $\alpha = 0$, we shall write, simply, βe. When we define addition and multiplication by the real number a for the expressions in question by the equations

$$(\alpha + \beta e) + (\gamma + \delta e) = (\alpha + \gamma) + (\beta + \delta)e, \tag{9}$$

$$(\alpha + \beta e)a = (\alpha a) + (\beta a)e, \tag{10}$$

we obtain an eight-dimensional real vector space with the basis

$$1, \ i, \ j, \ k, \ e, \ ie, \ je, \ ke. \tag{11}$$

In this space we define a multiplication by the equation

$$(\alpha+\beta e)(\gamma+\delta e)=(\alpha\gamma-\bar{\delta}\beta)+(\delta\alpha+\beta\bar{\gamma})e. \tag{12}$$

It is easy to verify the distributivity of this multiplication with respect to the addition (9) and also, by virtue of (10), the validity of the equations

$$[(\alpha+\beta e)(\gamma+\delta e)]a=[(\alpha+\beta e)a](\gamma+\delta e)=(\alpha+\beta e)[(\gamma+\delta e)a].$$

The multiplication table of this Cayley algebra in the basis (11) is easily derived from (12). It is as follows:

	1	i	j	k	e	ie	je	ke
1	1	i	j	k	e	ie	je	ke
i	i	-1	k	$-j$	ie	$-e$	$-ke$	je
j	j	$-k$	-1	i	je	ke	$-e$	$-ie$
k	k	j	$-i$	-1	ke	$-je$	ie	$-e$
e	e	$-ie$	$-je$	$-ke$	-1	i	j	k
ie	ie	e	$-ke$	je	$-i$	-1	$-k$	j
je	je	ke	e	$-ie$	$-j$	k	-1	$-i$
ke	ke	$-je$	ie	e	$-k$	$-j$	i	-1

We see that the elements of the form $\alpha=\alpha+0e$ constitute a subalgebra of the Cayley algebra isomorphic to the quaternion algebra. Furthermore, it follows from (9) and (12) that both the addition and the multiplication in the expression $\alpha+\beta e$ for the elements of the Cayley algebra can be interpreted in the sense of the operations as they are defined in the algebra.

The Cayley algebra is neither commutative nor associative. For example,

$$(ij)e=ke, \qquad i(je)=-ke.$$

11. If $\xi=\alpha+\beta e$ is an element of the Cayley algebra, we call the element

$$\bar{\xi}=\bar{\alpha}-\beta e$$

the *conjugate* of ξ. On the basis of (12) and (5), it is easy to verify that for arbitrary elements ξ and η of the Cayley algebra, we have

$$\bar{\xi}\bar{\eta}=\bar{\eta}\bar{\xi}, \quad \overline{\xi+\eta}=\bar{\xi}+\bar{\eta} \tag{13}$$

and that for $\xi=\alpha+\beta e$, we have

$$\xi\bar{\xi}=\bar{\xi}\xi=\alpha\bar{\alpha}+\beta\bar{\beta}=n(\alpha)+n(\beta). \tag{14}$$

This non-negative real number, which is equal to zero for $\xi = 0$ only, is called the *norm* of ξ and is denoted by $n(\xi)$.

If $\xi = \alpha + \beta e$ and $\eta = \gamma + \delta e$, then it follows from (12) and (14), in virtue of (5), that

$$n(\xi\eta) = (\alpha\gamma - \bar{\delta}\beta)(\bar{\gamma}\bar{\alpha} - \bar{\beta}\delta) + (\delta\alpha + \beta\bar{\gamma})(\bar{\alpha}\bar{\delta} + \gamma\bar{\beta}) =$$
$$= n(\alpha)n(\gamma) + n(\beta)n(\delta) + n(\alpha)n(\delta) + n(\beta)n(\gamma) + a - b,$$

where

$$a = \delta\alpha\gamma\bar{\beta} + \beta\bar{\gamma}\bar{\alpha}\bar{\delta},$$
$$b = \alpha\gamma\bar{\beta}\delta + \bar{\delta}\beta\bar{\gamma}\bar{\alpha}.$$

Since by (5) the second terms in the expressions for a and b are conjugate to the first terms, a and b are real numbers. If $\delta = 0$, then obviously $a = b = 0$, and so $a - b = 0$. But if $\delta \neq 0$, so that $n(\delta) \neq 0$, then, since a is a real number,

$$an(\delta) = \bar{\delta}\alpha\delta = bn(\delta),$$

and hence $a = b$, so that, again, $a - b = 0$.

Thus,

$$n(\xi\eta) = [n(\alpha) + n(\beta)][n(\gamma) + n(\delta)] = n(\xi)n(\eta). \tag{15}$$

Hence it follows that *the Cayley algebra has no divisors of zero.*

12. The equation (12) shows that the element $1 = 1 + 0e$ is the unit element of the Cayley algebra.

Furthermore, if $\xi = \alpha + \beta e$ and $\eta = \gamma + \delta e$ are any two elements with $\xi \neq 0$, then the equation

$$\xi\zeta = \eta \tag{16}$$

has a solution; and, because there are no divisors of zero, this solution is unique:

$$\zeta = [n(\xi)]^{-1}\bar{\xi}\eta = [n(\xi)]^{-1}[(\bar{\alpha}\gamma + \bar{\delta}\beta) + (\delta\bar{\alpha} - \beta\bar{\gamma})e)].$$

The reader can easily verify, using (12) and (5), that the element ζ does, in fact, satisfy the equation (16). Similarly, the unique solution of the equation

$$\zeta\xi = \eta \qquad (\xi \neq 0)$$

is the element

$$\zeta = [n(\xi)]^{-1}\eta\bar{\xi}$$

Thus, we see that *the Cayley algebra is a (non-associative) field.*

§ 37. Alternative Rings. Artin's Theorem

1. The quaternion algebra and the Cayley algebra should not be regarded as purely incidental examples of real division algebras: their fundamental role becomes apparent in many theorems, some of which will be proved, or at least mentioned, in the next section and also in §§ 46 and 50.

First of all, let us indicate a simple classification of all rings. The associative law is, after all, a relation among three elements, so that *every ring in which all the subrings generated by three elements are associative is itself associative.* A wider class of rings is constituted by the *alternative rings,* that is, rings in which all the subrings generated by *two* elements are associative. A still wider class are the *power-associative rings,* that is, rings in which all the subrings generated by a single element are associative; this class does not, of course, exhaust the class of all possible rings.

2. If a, b, c are elements of a ring R, then we define the *associator* of these elements as

$$[a, b, c] = (ab)c - a(bc). \tag{1}$$

Obviously,

$$[a, b, c] = 0$$

if and only if the elements a, b, c satisfy the associative law

$$(ab)c = a(bc).$$

From (1) it follows that

$$[a + a', b, c] = [a, b, c] + [a', b, c]; \tag{2}$$

similar equations are valid for the places occupied in the associator by b and c. Furthermore,

$$[-a, b, c] = -[a, b, c].$$

It is easy to verify the following equation, which holds for arbitrary elements $a, b, c, d \in R$:

$$[ab, c, d] - [a, bc, d] + [a, b, cd] = a[b, c, d] + [a, b, c]d. \tag{3}$$

3. ARTIN'S THEOREM: *A ring R is alternative if and only if the equations*

$$(aa)b = a(ab), \quad (ba)a = b(aa) \tag{4}$$

hold[1] *for arbitrary $a, b \in R$.*

[1] If the reader wishes, he may without harm take the identity (4) as the definition of an alternative ring and omit the proof of Artin's Theorem. We see that alternative rings form a primitive class of universal algebras (see § 22.3).

It is obvious that in an alternative ring the conditions (4) are satisfied. We shall therefore assume that R is a ring in which (4) holds for arbitrary elements a and b; these equations can be written in the form

$$[a, a, b] = 0, \qquad [b, a, a] = 0. \tag{5}$$

4. *Arbitrary elements a and b of R satisfy the equation*

$$[a, b, a] = 0. \tag{6}$$

For by (5) and (2),

$$0 = [a, a + b, a + b] =$$
$$= [a, a, a] + [a, a, b] + [a, b, a] + [a, b, b] = [a, b, a].$$

If the elements a, b, c of R are permuted, then the associator $[a, b, c]$ remains unchanged when the permutation is even and changes sign when it is odd.

It is, of course, sufficient to discuss the case of a transposition of two elements. By (5) and (2),

$$0 = [a, b + c, b + c] =$$
$$= [a, b, b] + [a, b, c] + [a, c, b] + [a, c, c] = [a, b, c] + [a, c, b].$$

Hence

$$[a, b, c] = -[a, c, b].$$

Similarly, we can show that

$$[a, b, c] = -[b, a, c]$$

and by (6) that

$$[a, b, c] = -[c, b, a].$$

5. If A, B, C are subsets of our ring R, then we shall write

$$[A, B, C] = O,$$

if $[a, b, c] = 0$ for all $a \in A$, $b \in B$, $c \in C$.

We call a subset A of R an *a-set* if

$$[A, A, R] = O, \tag{7}$$

or, by what has been proved above, if the equivalent conditions

$$[A, R, A] = O, \qquad [R, A, A] = O \tag{8}$$

hold. From (5) it follows that a set consisting of a single element is an α-set.

The subring of R generated by an α-set A is itself an α-set; for brevity, we call it an *α-subring.*

For it follows from (2) that by adjoining to A all possible sums and differences of its elements we again obtain an α-set. Let us show that this is also the case when we add to A the product of all possible pairs of its elements. In fact, if $a_1, a_2, a_3, a_4 \in A$, then by (3),

$$[a_1a_2, x, a_3] - [a_1, a_2x, a_3] + [a_1, a_2, xa_3] = \\ = a_1[a_2, x, a_3] + [a_1, a_2, x]a_3.$$

Hence by (7) and (8),

$$[a_1a_2, x, a_3] = 0. \tag{9}$$

Further, again by (3),

$$[a_1a_2, x, a_3a_4] - [a_1, a_2x, a_3a_4] + [a_1, a_2, x(a_3a_4)] = \\ = a_1[a_2, x, a_3a_4] + [a_1, a_2, x](a_3a_4);$$

hence by (7), (8), and (9),

$$[a_1a_2, x, a_3a_4] = 0.$$

Thus, if we begin with A and alternately add to an α-set already constructed all the sums and differences of the elements of the α-set and then all the products of pairs of elements, we obtain an ascending sequence of α-sets of R. The union of this sequence is, of course, an α-set and coincides with the subring generated by A.

From this result it follows that *every subring of R generated by a single element is an α-set and is therefore associative.*

6. If A and B are α-subrings of R, then we denote by C the set of all those elements c of R for which

$$[A, B, c] = O.$$

Clearly,

$$C \supseteq (A \cup B), \tag{10}$$

and C is closed with respect to addition and subtraction.

If $c \in C$, then for arbitrary $a' \in A$ and $b' \in B$ we have

$$a'c, ca', b'c, cb' \in C.$$

For if $a \in A$, $b \in B$, then by (3),

$$[aa', c, b] - [a, a'c, b] + [a, a', cb] = a[a', c, b] + [a, a', c]b,$$

$$[bc, a', a] - [b, ca', a] + [b, c, a'a] = b[c, a', a] + [b, c, a']a,$$

or, bearing in mind that A is an α-set and $c \in C$,

$$[a, a'c, b] = 0,$$
$$[b, ca', a] = 0,$$

i.e., $a'c \in C$, $ca' \in C$. The other two relations are proved in the same way.

By § 9.7, every element of the subring $\{A, B\}$ can be represented as a sum of products of the form $w = x_1 x_2 \dots x_n$, where every x_i $(i = 1, 2, \dots, n)$ belongs to A or to B; we call n the *length* of this product: $n = l(w)$. Parentheses are distributed in some way throughout w, and every parenthesis is the product of exactly two lesser parentheses. A product w of length n is called a *normal element* if every parenthesis of length k $(2 \leqq k \leqq n)$ that occurs in the formation of this product is a product of a parenthesis of length $k-1$ by an element of A or of B, on the left or on the right. For example, the product $(ab)(b'a')$ is not normal, but the product $b'[(ab)a']$ is.

From what has been proved above, it follows by (10) that every normal element w belongs to C, that is, for all $a \in A$, $b \in B$,

$$[a, b, w] = 0. \tag{11}$$

7. *Every product of normal elements can be represented as a sum of normal elements.*

It is sufficient to prove this for a product of two normal elements v and w. The proof will be by induction on the length of v. To begin with, it is obvious that vw is normal for $l(v) = 1$.

If $v = v'a$, $l(v') = l(v) - 1$, $a \in A$, then

$$vw = (v'a)w = v'(aw) + [v', a, w] =$$
$$= v'(aw) - [v', w, a] = v'(aw) - (v'w)a + v'(wa);$$

but, by the induction hypothesis, each of the three terms of the sum can be represented as a sum of normal elements. The case $v = av'$ can be reduced to the following:

$$vw = (av')w = a(v'w) + [a, v', w] =$$
$$= a(v'w) - (v', a, w) = a(v'w) - (v'a)w + v'(aw).$$

Hence it follows that *every element of* $\{A, B\}$ *can be represented as a sum of normal elements.*

8. *If v and w are normal elements, then for arbitrary $a \in A$, $b \in B$,*

$$[a, v, w] = [b, v, w] = 0. \tag{12}$$

If $l(v)=1$, then (12) follows from (11). We shall therefore give a proof by induction on the length of v. If $v=b'v'$, where $l(v')=l(v)-1$, $b'\in B$, then by (3),

$$[b'v', w, a]-[b', v'w, a]+[b', v', wa]=b'[v', w, a]+[b', v', w]a.$$

Hence, by using the induction hypothesis, the formula (11), and the result proved in the preceding paragraph, we obtain

$$[b'v', w, a]=[v, w, a]=[a, v, w]=0.$$

In the cases $v=v'a'$, $v=a'v'$, and $v=v'b'$, the identity (3) is also applicable, but for our a, b, c, d we must take w, v', a', a in the first case, a, a', v', w in the second case, and v', b', w, a in the third.

9. *If u, v, and w are normal elements, then*

$$[u, v, w]=0. \tag{13}$$

By (12), we can conduct the proof by induction on the length of the element u. If $u=u'a$, $l(u')=l(u)-1$, $a\in A$, then by (3),

$$[u'a, v, w]-[u', av, w]+[u', a, vw]=u'[a, v, w]+[u', a, v]w,$$

or, by using the induction hypothesis and the result of § 37.7,

$$[u'a, v, w]=[u, v, w]=0.$$

The other possible cases are discussed in the same way.

Taking into account that every element of the subring $\{A, B\}$ is a sum of normal elements, we now obtain that *the subring $\{A, B\}$ generated by the α-subrings A and B is associative.* If we take for A and B subrings each of which is generated by a single element, then we see, in virtue of the last result of § 37.5, that *every subring of R generated by two elements is associative,* and this concludes the proof of Artin's Theorem. //

❖ If the additive group of a ring R does not contain non-zero elements of finite order, then R is a power-associative ring if and only if for every element a of R

$$(aa)a=a(aa),$$
$$[(aa)a]a=(aa)(aa).$$

[A. A. Albert, *Summa Bras. Math.*, vol. 2 (1948), pp. 21-33; see also A. T. Gainov, *Usp. Mat. Nauk*, vol. 12 (1957), pp. 141-46]. ❖

10. We shall now show that *the Cayley algebra is alternative.*

By Artin's Theorem it is sufficient to verify that the equations (4) hold in the Cayley algebra. We shall only verify the first of these equations. If

$$\xi = \alpha + \beta e, \quad \eta = \gamma + \delta e,$$

then, by § 36.10 (12),

$$(\xi\xi)\eta = [(\alpha^2 - \bar{\beta}\beta)\gamma - \bar{\delta}(\beta\alpha + \beta\bar{\alpha})] + [\delta(\alpha^2 - \bar{\beta}\beta) + (\beta\alpha + \beta\bar{\alpha})\bar{\gamma}]e,$$
$$\xi(\xi\eta) = [\alpha(\alpha\gamma - \bar{\delta}\beta) - \overline{(\delta\alpha + \beta\bar{\gamma})}\beta] + [(\delta\alpha + \beta\bar{\gamma})\alpha + \beta\overline{(\alpha\gamma - \bar{\delta}\beta)}]e.$$

But the right-hand sides of these two equations coincide, as is easy to show on the basis of (5) of § 36.9 if we take into account that both $\alpha + \bar{\alpha}$ and $\bar{\beta}\beta = \beta\bar{\beta}$ are real numbers and are therefore permutable with arbitrary quaternions. //

§ 38. The Generalized Frobenius Theorem

1. *The field of real numbers and the field of complex numbers are the only finite-dimensional real associative-commutative algebras without divisors of zero.*

The field of quaternions is the only finite-dimensional real associative, but not commutative algebra without divisors of zero.

The Cayley algebra is the only finite-dimensional real alternative, but not associative algebra without divisors of zero.

The first two theorems combined are called *Frobenius' Theorem*; we shall call all three together, the *Generalized Frobenius Theorem.*

2. To begin with, we prove the following general theorem:

Every finite-dimensional algebra R without divisors of zero over an arbitrary commutative field P is an algebra with unique division.

We shall only prove that the first of the equations

$$ax = b, \qquad ya = b, \tag{1}$$

where $a \neq 0$, has a unique solution. We denote by aR the set of elements of the form ar for all $r \in R$. It is clear that aR is a linear subspace of the additive vector space of the algebra R. If this subspace is not the whole algebra R, then it has a strictly smaller dimension. Therefore, by taking an arbitrary basis

$$x_1, x_2, \ldots, x_n$$

of R we see that the elements

$$ax_1, ax_2, \ldots, ax_n$$

are linearly dependent, i.e., that the field P contains elements $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all equal to zero, such that

$$(ax_1)\alpha_1 + (ax_2)\alpha_2 + \ldots + (ax_n)\alpha_n = 0.$$

Hence it follows, however, that

$$a(x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n) = 0,$$

where both factors are different from zero, in contradiction to the condition that R has no divisors of zero.

Thus, we see that $aR = R$. There exists an element $r \in R$, therefore, for which

$$ar = b,$$

so that the first of the equations (1) has, in fact, proved to be solvable. The uniqueness of the solution follows from the absence of divisors of zero in R. //

For later use, we mention that in an alternative algebra not only is every subring generated by two elements associative, but so is every subalgebra with two generators: we only have to bear in mind how the elements of this subalgebra are expressed in terms of the elements of the subring generated by these two generators.

Every alternative finite-dimensional algebra with unique division R over an arbitrary commutative field P has a unit element.

In the associative case, the statement of this theorem follows immediately from § 8.10. But if R is merely an alternative algebra, then each of its subalgebras with two generators is an associative finite-dimensional algebra without divisors of zero, so that, by the preceding theorem, it is an algebra with unique division and therefore has a unit element. Suppose now that two such subalgebras have distinct unit elements e_1 and e_2. The subalgebra generated by these two elements is itself an associative division algebra, and we arrive at a contradiction, because an associative field cannot contain two distinct non-zero *idempotent elements,* that is, elements equal to their squares. //

3. Proceeding now to the proof of the Generalized Frobenius Theorem, we may assume that R is a real alternative algebra (in particular, an associative, or even associative-commutative algebra) with unique division, of finite dimension n, and having a unit element.

As was shown in § 36.2, the center of R contains a subfield D isomorphic to the field of real numbers: this is the set of elements that are multiples of the unit element 1 of R.

If an element a of R does not belong to the subfield D, then the subalgebra $\{a\}$ generated by a contains this subfield and is isomorphic to the field of complex numbers.

For since R is an alternative algebra, $\{a\}$ is associative, so that we can speak of the powers of a in the usual sense. Since R is of dimension n, the elements

$$1, a, a^2, \ldots, a^n$$

are linearly dependent. Therefore, there exists a polynomial $f(x)$ with real coefficients, of degree not exceeding n, that vanishes for a. From the Fundamental Theorem of Algebra it follows that $f(x)$ factors into the product of linear polynomials and irreducible quadratic polynomials both with real coefficients; and since R has no divisors of zero, at least one of these factors vanishes for a; we denote it by $\varphi(x)$.

Since a does not lie in D and therefore does not satisfy any equation of degree 1 with real coefficients, we have

$$\varphi(x) = x^2 + \beta x + \gamma \qquad (\gamma \neq 0). \tag{2}$$

Thus,

$$a^2 + \beta a + \gamma = 0, \tag{3}$$

so that $\gamma \in \{a\}$, and hence the whole of D is contained in $\{a\}$. From (3) it follows that $\{a\}$ is two-dimensional and that the elements 1 and a, where

$$a^2 = -\gamma - \beta a.$$

form a basis for it.

Next, we denote by α a complex number that is a root of the quadratic trinomial (2); α is not real, so that the numbers 1 and α form a basis of the algebra of complex numbers. However, the multiplication table defining the algebra of complex numbers in this basis is obviously the same as that defining $\{a\}$ in the basis $1, a$. Hence, by § 36.3, it follows that the subalgebra $\{a\}$ is isomorphic to the field of complex numbers.

4. *If the elements a and b of R are not in D and if they generate distinct subalgebras $\{a\}$, $\{b\}$, then the subalgebra $\{a, b\}$ is isomorphic to the quaternion algebra.*

Since $\{a\}$ is isomorphic to the field of complex numbers, we can choose in it a basis $1, i$, where

$$i^2 = -1. \tag{4}$$

Similarly, $\{b\}$ has a basis $1, j_0$, with

$$j_0^2 = -1. \tag{5}$$

Clearly, $\{i\} = \{a\}$ and $\{j_0\} = \{b\}$, and hence $\{i, j_0\} = \{a, b\}$; and since $j_0 \notin a$, the elements $1, i, j_0$ are linearly independent.

Hence both the elements $i + j_0, i - j_0$ are not in D and each must therefore be the root of an irreducible quadratic polynomial with real coefficients; but then the squares of these elements must be expressible in terms of their first powers and the unit element. There exist, then, real numbers $\alpha, \beta, \gamma, \delta$ such that

$$\begin{aligned}(i + j_0)^2 &= -2 + (ij_0 + j_0 i) = \alpha(i + j_0) + \beta, \\ (i - j_0)^2 &= -2 - (ij_0 + j_0 i) = \gamma(i - j_0) + \delta.\end{aligned} \tag{6}$$

Addition of these equations gives

$$-4 = (\alpha + \gamma)i + (\alpha - \gamma)j_0 + (\beta + \delta).$$

And hence, by the linear independence of the elements $1, i, j_0$, it follows that $\alpha + \gamma = \alpha - \gamma = 0$, so that $\alpha = \gamma = 0$. From (6) we now find that $ij_0 + j_0 i$ is a real number, which we denote by 2μ: we have

$$ij_0 + j_0 i = 2\mu = \beta + 2 = -(\delta + 2). \tag{7}$$

Since the quadratic trinomials that the elements $i + j_0$ and $i - j_0$ satisfy are irreducible, we must have $\beta < 0, \delta < 0$ on account of $\alpha = \gamma = 0$, so that by (7),

$$-1 < \mu < 1.$$

The number

$$\nu = 1/\sqrt{1 - \mu^2} \tag{8}$$

is therefore real and distinct from zero.

We now consider the element

$$j = \mu\nu i + \nu j_0$$

of the subalgebra $\{a, b\}$. Since $\nu \neq 0$, the elements $1, i, j$ are linearly independent, and on the basis of (7) and (8) it is easy to verify the equations

$$j^2 = -1, \tag{9}$$

$$ij + ji = 0. \tag{10}$$

We introduce the notation

$$k = ij = -ji. \tag{11}$$

If we had an equation

$$k = \alpha + \beta i + \gamma j$$

with real coefficients α, β, γ, then by multiplying both sides on the right by i we would obtain, by (11),

$$j = \alpha i - \beta - \gamma k = \alpha i - \beta - \gamma(\alpha + \beta i + \gamma j).$$

However, by comparing the coefficients of j we would obtain the equation

$$1 = -\gamma^2,$$

which is impossible, because γ is a real number.

Thus, the elements $1, i, j, k$ have proved to be linearly independent. Since R is an alternative algebra, these elements lie in the associative subalgebra $\{a, b\}$ and even generate this subalgebra. Therefore, by using (4), (9), and (11), it is not difficult to verify that the elements $1, i, j, k$ satisfy all the equations (4) of § 36.8. For example,

$$k^2 = (ij)(-ji) = -i(j^2)i = i^2 = -1,$$
$$jk = j(-ji) = -(j^2)i = i.$$

This shows that the elements $1, i, j, k$ form a basis of the algebra $\{a, b\}$ generated by them and that this subalgebra of R is isomorphic to the quaternion algebra.

Thus, the first theorem of § 38.1 is proved. //

5. If the proof of the second theorem of § 38.1 were our ultimate aim, there would be very little left to do. Suppose that the algebra R in question is associative and that it contains a subalgebra $\{a, b\}$ isomorphic to the quaternion algebra and an element c not in $\{a, b\}$. We denote by $1, i, j, k$ a basis of $\{a, b\}$ with the usual multiplication table for the quaternion algebra: such a basis we shall call *canonical*. Furthermore, under our assumptions, the subalgebra $\{c\}$ is isomorphic to the field of complex numbers and therefore contains an element e for which

$$e^2 = -1,$$

with $\{e\} = \{c\}$.

Repeating the arguments used in the derivation of (7), we can claim that there exist real numbers α, β, γ such that

$$ie + ei = \alpha, \quad je + ej = \beta, \quad ke + ek = \gamma.$$

As R is an associative algebra, we now obtain

$$ek = (ei)j = \alpha j - i(ej) = \alpha j - \beta i + (ij)e = \alpha j - \beta i + ke =$$
$$= \alpha j - \beta i + \gamma - ek,$$

i.e.,

$$2ek = \alpha j - \beta i + \gamma,$$

or, after multiplication on the right by k,

$$-2e = \alpha i + \beta j + \gamma k.$$

But this leads to the relation $e \in \{a, b\}$, which is impossible. And so the second theorem of § 38.1 is proved. //

6. We now turn to the proof of the Generalized Frobenius Theorem.

If the elements a, b, and c of R are such that $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$ are distinct subalgebras isomorphic to the field of quaternions, then the subalgebra $\{a, b, c\}$ is isomorphic to the Cayley algebra.

Since the subalgebra $\{a, b\}$ is isomorphic to the field of quaternions, it has a canonical basis $1, i, j, k$. On the other hand, the subalgebra $\{c\}$ contains an element e_0 for which

$$e_0^2 = -1,$$

with $\{e_0\} = \{c\}$.

The subalgebra $\{i, e_0\}$ is isomorphic to the field of quaternions, because it would follow from $\{i\} = \{e_0\}$ that $c \in \{a, b\}$. Therefore, as proved in § 38.4, there exists an element

$$e_1 = \alpha_1 i + \beta_1 e_0 \tag{12}$$

with real α_1 and β_1 ($\beta_1 \neq 0$) such that the elements $1, i, e_1, ie_1$ form a canonical basis of $\{i, e_0\}$.

From $\{j\} = \{e_1\}$ it would follow by (12) that $e_0 \in \{a, b\}$, which is impossible. Therefore $\{j, e_1\}$ is isomorphic to the field of quaternions and hence contains an element

$$e_2 = \alpha_2 j + \beta_2 e_1 \tag{13}$$

with real α_2 and β_2 ($\beta_2 \neq 0$) such that the elements $1, j, e_2, je_2$ form a canonical basis of $\{j, e_2\}$.

We shall now show that *the subalgebra $\{i, e_2\}$ is also isomorphic to the field of quaternions and has the system of elements $1, i, e_2, ie_2$ as a canonical basis.* In fact, from $\{i\} = \{e_2\}$ it would follow by (13) and (12) that $e_1 \in \{a, b\}$ and hence that $e_0 \in \{a, b\}$, which is impossible. Thus, we have proved the first statement. Next, since it is already known that

$$i^2 = e_2^2 = -1,$$

it is sufficient for the proof of the second statement, as we know from § 38.4, to establish the equation

$$ie_2 + e_2 i = 0.$$

However, bearing in mind that the equations

$$ij + ji = 0, \quad ie_1 + e_1 i = 0$$

hold, we obtain

$$ie_2 + e_2 i = i(\alpha_2 j + \beta_2 e_1) + (\alpha_2 j + \beta_2 e_1)i = \alpha_2(ij + ji) + \beta_2(ie_1 + e_1 i) = 0.$$

7. As above, the subalgebra $\{k, e_2\}$ is isomorphic to the field of quaternions and has a canonical basis $1, k, e, ke$, where

$$e = \alpha_3 k + \beta_3 e_2 \tag{14}$$

with real α_3 and β_3 ($\beta_3 \neq 0$). Note that $\{k, e_2\} = \{k, e\}$. As in the preceding subsections, it can be verified that the subalgebras $\{i, e\}$ and $\{j, e\}$ are also isomorphic to the field of quaternions and have as canonical bases the systems of elements $1, i, e, ie$ and $1, j, e, je$, respectively.

8. *The system of elements*

$$1,\ i,\ j,\ k,\ e,\ ie,\ je,\ ke \tag{15}$$

is linearly independent.

For if it were linearly dependent, then the subalgebra $\{i, j\} = \{a, b\}$ would contain elements u, v, with $v \neq 0$, such that

$$u = ve. \tag{16}$$

Since the subalgebra $\{v\}$ is isomorphic to the field of complex numbers or that of real numbers, the element v^{-1} is contained in it and therefore in the associative subalgebra $\{v, e\}$. Hence

$$v^{-1}(ve) = (v^{-1}v)e = e,$$

i.e., from (16) we would obtain that

$$e = v^{-1}u \in \{a, b\},$$

which is impossible.

9. *The elements* (15) *satisfy the multiplication table of the Cayley algebra* (see § 36.10).

We know that the systems of elements

$$(1, i, j, k),\ (1, i, e, ie),\ (1, j, e, je),\ (1, k, e, ke) \tag{17}$$

are canonical bases of the subalgebras isomorphic to the quaternion algebra generated by each of them. Because of this, there are only a few entries

in this multiplication table to be verified. We choose a few typical cases to show how this can be done.

Since the subalgebra $\{e+i, j\}$ is associative, we have

$$[j(e+i)](e+i) = j[(e+i)(e+i)],$$

and hence, by using what we have said above about the systems (17), we obtain

$$-j+(je)i-ke-j=-2j,$$

or

$$(je)i = ke. \tag{18}$$

Next, from the equation

$$(e+i)[j(e+i)] = [(e+i)j](e+i)$$

we obtain with the help of (18) that

$$i(je) = -ke. \tag{19}$$

Finally, since the subalgebra $\{k, i+je\}$ is associative, we have the equation

$$[(i+je)(i+je)]k = (i+je)[(i+je)k],$$

from which, with the help of (18), (19) and the equation

$$(je)k = -ie,$$

which can be proved similarly, we obtain

$$(je)(ie) = k.$$

In this way it can be shown that the elements (15) form a basis of the subalgebra $\{a, b, c\}$ generated by them and that this subalgebra is isomorphic to the Cayley algebra.

A basis of a Cayley subalgebra satisfying the multiplication table of § 36.10 will be called *canonical*.

10. Finally, let us assume that the algebra R contains a subalgebra isomorphic to the Cayley algebra and having as a canonical basis the system of elements

$$1, i, j, k, e, ie, je, ke,$$

as well as some element not in this subalgebra. Then, by generalizing the discussion given in the preceding subsections, we could find an element f

in R such that each of the following systems of elements would generate a subalgebra isomorphic to the Cayley algebra and be a canonical basis of it:

$$1,\ i,\ \ j,\ k,\ \ f,\ if,\ \ \ jf,\ kf;$$
$$1,\ i,\ \ e,\ ie,\ \ f,\ if,\ \ \ ef,\ (ie)f;$$
$$1,\ j,\ \ e,\ je,\ \ f,\ jf,\ \ \ ef,\ (je)f;$$
$$1,\ je,\ i,\ ke,\ f,\ (je)f,\ if,\ (ke)f.$$

The corresponding constructions are very tedious, but they do not present any fundamental difficulty and can therefore be left to the reader.

Since these bases are canonical, we derive the following equations:

$$e^2 = -1, \qquad (if)^2 = -1,$$
$$(jf)(if) = k, \qquad k(if) = -jf,$$
$$e(if) = (ie)f, \qquad (if)e = -(ie)f,$$
$$(jf)e = -(je)f, \qquad [(je)f]e = jf,$$
$$[(je)f](if) = -ke.$$

By use of these equations we obtain from the equation

$$(jf)[(e+if)(e+if)] = [(jf)(e+if)](e+if),$$

which holds because R is an alternative algebra, that

$$-2(jf) = [-(je)f + k](e+if) = -2(jf) + 2ke,$$

i.e., $ke = 0$, which is impossible.

This contradiction concludes the proof of the Generalized Frobenius Theorem. //

❖ Every alternative field is either associative or, regarded as an algebra over its center (see § 36.2), is finite-dimensional—specifically, of dimension 8—and is the so-called Cayley-Dickson algebra, that is, a generalization of the Cayley algebra to the case of an arbitrary ground field [L. A. Skornjakov, *Ukr. Mat. Z.*, vol. 2 (1950), pp. 70-85; R. H. Bruck and E. Kleinfeld, *Proc. Amer. Math. Soc.*, vol. 2 (1951), pp. 878-90]. ❖

11. The Frobenius Theorem cannot be generalized to the case of non-alternative algebras, since there exist, for example, very many distinct finite-dimensional real power-associative algebras with unique division.

❖ The dimension of a finite-dimensional real algebra without divisors of zero can only have the values $n = 1, 2, 4$, or 8 [J. Milnor, *Bull. Amer. Math. Soc.*, vol. 64 (1958), pp. 87-89]. ❖

12. Arguments paralleling those used in § 38.3 immediately lead to the following theorem.

The only finite-dimensional power-associative algebra over the field of complex numbers having a unit element but no divisors of zero is the field of complex numbers itself.

§ 39. The Birkhoff-Witt Theorem on Lie Algebras

1. In § 8.3 it was shown that the replacement of the operation of multiplication ab in an associative ring R by the operation of commutation

$$a \circ b = ab - ba,$$

makes it into a Lie ring $L(R)$.

If R is a Σ-operator ring (see § 31.7), *then $L(R)$ is also a Σ-operator ring.*

Since these rings have the same additive group, we have only to verify the validity of the condition (11) of § 31.7. If $a, b \in R$, $\alpha \in \Sigma$, then

$$(a \circ b)\alpha = (ab - ba)\alpha = (ab)\alpha - (ba)\alpha = (a\alpha)b - b(a\alpha) = (a\alpha) \circ b.$$

In the same way we can verify the equation

$$(a \circ b)\alpha = a \circ (b\alpha).$$

In particular, *if R is an algebra over a commutative field P, then $L(R)$ is also an algebra over the same field.*

2. In the case of algebras, the results of § 8.3 have a certain converse in the form of the following theorem [G. D. Birkhoff, *Ann. of Math.*, vol. 38 (1937), pp. 526-32; E. Witt, *J. reine und angew. Math.*, vol. 177 (1937), pp. 152-60].

For every Lie algebra L over an arbitrary commutative field P there exists an associative algebra R over the same field such that L can be isomorphically embedded in the algebra $L(R)$.

Proof. We choose a basis of L over P. In accordance with Zermelo's Theorem (see § 6.3), we may assume that this basis is well-ordered, and we write it in the form

$$e_1, e_2, \ldots, e_\alpha, \ldots. \tag{1}$$

If we denote the multiplication in L by the symbol $\times$, then

$$e_\alpha \times e_\beta = \sum_\gamma c_{\alpha\beta}^\gamma e_\gamma; \tag{2}$$

clearly, for given α and β only a finite number of the coefficients $c_{\alpha\beta}^\gamma$ can be different from zero.

We define a *word* as an ordered finite system of elements of the basis (1), not necessarily distinct. Let

$$w = e_{\alpha_1} e_{\alpha_2} \dots e_{\alpha_k} \tag{3}$$

be a word of *length* k ($k \geqq 1$): we write the elements forming the word next to each other, without separation by commas. As usual, every pair of elements e_{α_i}, e_{α_j} of the word for which $i < j$, but $\alpha_i > \alpha_j$ in the sense of the ordering of the basis (1) is called an *inversion*.

Furthermore, we shall consider *sums of words*, that is, finite unordered systems of words, not necessarily distinct, each having a non-zero coefficient from the field P. In writing down sums of words we shall connect the words formally by the symbol $+$.

The greatest length of the words that occur in a given sum of words s is called the *degree* of this sum. If n is the degree of s, then we denote by l_i ($i = n, n-1, \dots, 2$) the total number of inversions in the words of length i occurring in s; if there are no words of length i in s, then we put $l_i = 0$. Of course, l_1 has no meaning. The symbol

$$\sigma = (l_n, l_{n-1}, \dots, l_2) \tag{4}$$

is called the *height* of s.

3. Taking n fixed, we introduce a lexicographical ordering in the set of heights of the form (4), i.e., of degree n; let

$$\sigma' = (l'_n, l'_{n-1}, \dots, l'_2);$$

then if there is an i such that $l_i < l'_i$ but (for $i < n$) $l_n = l'_n, \dots, l_{i+1} = l'_{i+1}$, we shall put $\sigma < \sigma'$.

The set of heights of degree n is a well-ordered set under lexicographic ordering.

Obviously, we have a linear ordering. Let us consider a strictly decreasing sequence of heights

$$\sigma_1 > \sigma_2 > \dots > \sigma_k > \dots, \tag{5}$$

where

$$\sigma_k = (l_{kn}, l_{k,n-1}, \dots, l_{k2}) \qquad (k = 1, 2, \dots).$$

Clearly,

$$l_{1n} \geqq l_{2n} \geqq \dots \geqq l_{kn} \geqq \dots,$$

and hence there exists a k_1 such that

$$l_{k_1 n} = l_{k_1+1,n} = \dots.$$

From this it follows that

$$l_{k_1,n-1} \geqq l_{k_1+1,n-1} \geqq \dots,$$

and so there exists a k_2 ($k_2 \geqq k_1$) such that

$$l_{k_2,n-1} = l_{k_2+1,n-1} = \dots.$$

Continuing further, for $n-2, \dots, 2$, we see that the sequence (5) terminates.

4. Suppose that the word $w = w_1 e_\beta e_\alpha w_2$ occurs in a sum of words s with the coefficient $a \in P$, where $\beta > \alpha$ in the sense of the ordering of the basis (1), and where w_1 and w_2 are words, one or both of which may be absent. We define a *reduction* of s as a replacement of the term aw in s by the sum of words

$$aw_1 e_\alpha e_\beta w_2 + \sum_\gamma (ac_{\beta\alpha}^\gamma) w_1 e_\gamma w_2.$$

A reduction obviously does not change the degree of s but reduces its height.

Hence it follows, in virtue of the well-ordering of the set of heights, that every sum of words s can be converted by a succession of reductions into a sum of height $(0, 0, \dots, 0)$, that is, a sum in which the summands are words without inversions or, as we shall call them, *normal words*. In the sum so obtained we then collect like terms; this may, of course, decrease the degree of the sum, but does not disturb its property of being a sum of normal words, and in this way we reduce s to *normal form*.

5. LEMMA: *Every sum of words can be reduced to a unique normal form, independent of the sequence of reductions performed.*

The lemma is true for sums of height $(0, 0, \dots, 0)$, because in this case all that is done is to collect like terms. This enables us to conduct the proof by an induction on the well-ordered set of heights.

Suppose that a sum of words s is reduced by two sequences of reductions to the normal forms s_1 and s_2, respectively. If both these chains begin with the same reduction carrying s into s', then the induction hypothesis applies, since the height of the sum s' is less than that of s; and therefore $s_1 = s_2$.

Suppose, on the other hand, that the initial reduction of the first chain and the initial reduction of the second concern distinct words of s and carry s into s' and s'', respectively. In s' we can perform the initial reduction of the second chain, thus arriving at a certain sum of words s''': we arrive at the same sum s''' by carrying out the initial reduction of the first chain in s''. If we now reduce s''' by any method to a normal form s_3 and bear in mind that the heights of the sums s' and s'' are less than that of s, we see that $s_1 = s_3$, $s_2 = s_3$, and hence $s_1 = s_2$.

The same arguments apply in the case in which the initial reductions of both chains refer to the two appropriate places of the word

$$w_1 e_\beta e_\alpha w_3 e_\delta e_\gamma w_2 \qquad (\beta > \alpha;\ \delta > \gamma), \tag{6}$$

whose coefficient we omit; the word w_3 may, of course, be absent. In this case, however, in order to go from the sums s' and s'' to the same sum s''' it is necessary to carry out the initial reduction of the second (first) chain in all those words of the sum s' (s'') that appear in place of (6).

6. It remains to consider the case where the initial reductions of both chains refer to the two appropriate places of the word

$$w_1 e_\gamma e_\beta e_\alpha w_2 \qquad (\gamma > \beta > \alpha). \tag{7}$$

As above, let s' and s'' be the sums of words obtained after these initial reductions. If the initial reduction of the first chain referred to the place (γ, β), then the word of the same length that is obtained in place of (7) as a result of this reduction is submitted to a further reduction at the place (γ, α) and then at the place (β, α). As a result we pass from s' to a sum s_1''' that differs from the original sum s by the fact that in place of (7) we have the sum

$$w_1 e_\alpha e_\beta e_\gamma w_2 + \sum_\delta c_{\gamma\beta}^\delta w_1 e_\delta e_\alpha w_2 + \sum_\delta c_{\gamma\alpha}^\delta w_1 e_\beta e_\delta w_2 + \sum_\delta c_{\beta\alpha}^\delta w_1 e_\delta e_\gamma w_2. \tag{8}$$

Similarly we can pass by a few reductions from s'' to a sum s_2''' that differs from s in the fact that (7) is replaced by the sum

$$w_1 e_\alpha e_\beta e_\gamma w_2 + \sum_\delta c_{\beta\alpha}^\delta w_1 e_\gamma e_\delta w_2 + \sum_\delta c_{\gamma\alpha}^\delta w_1 e_\delta e_\beta w_2 + \sum_\delta c_{\gamma\beta}^\delta w_1 e_\alpha e_\delta w_2. \tag{9}$$

7. Let us compare the sums (8) and (9). They contain the terms

$$c_{\gamma\beta}^\delta w_1 e_\delta e_\alpha w_2 \quad \text{and} \quad c_{\gamma\beta}^\delta w_1 e_\alpha e_\delta w_2. \tag{10}$$

If $\delta = \alpha$, then the terms (10) coincide. If $\delta > \alpha$, then a reduction is possible in the first of these terms that carries it into the second term to which the sum $\sum_\epsilon c_{\gamma\beta}^\delta c_{\delta\alpha}^\epsilon w_1 e_\epsilon w_2$ has been added. If, finally, $\alpha > \delta$, then a reduction in the second term of (10) carries it into the sum of the first term and of $\sum_\epsilon c_{\gamma\beta}^\delta c_{\alpha\delta}^\epsilon w_1 e_\epsilon w_2$.

Similar arguments apply to the other terms of (8) and (9). Hence it follows that these sums, after the appropriate reductions, can be transformed into sums—we denote them by (8′) and (9′)—that differ from each other only by words of the form $w_1 e_\epsilon w_2$. It is easy to see, however, that the sum of the coefficients in the given word $w_1 e_\epsilon w_2$ in (8′) is equal to

$$\sum_{\delta > \alpha} c_{\gamma\beta}^\delta c_{\delta\alpha}^\epsilon + \sum_{\delta < \beta} c_{\gamma\alpha}^\delta c_{\beta\delta}^\epsilon + \sum_{\delta > \gamma} c_{\beta\alpha}^\delta c_{\delta\gamma}^\epsilon \tag{11}$$

and in (9′) to

$$\sum_{\delta<\alpha} c_{\gamma\beta}^{\delta} c_{\alpha\delta}^{\epsilon} + \sum_{>\beta} c_{\gamma\alpha}^{\delta} c_{\delta\beta}^{\epsilon} + \sum_{\delta<\gamma} c_{\beta\alpha}^{\delta} c_{\gamma\delta}^{\epsilon}. \tag{12}$$

These two sums of coefficients are equal. For when we apply the multiplication table (2) to the equation

$$(e_\gamma \times e_\beta) \times e_\alpha + (e_\beta \times e_\alpha) \times e_\gamma + (e_\alpha \times e_\gamma) \times e_\beta = 0,$$

which holds in L, and equate the coefficient of e_ϵ to zero, we obtain the equation

$$\sum_{\delta} \left(c_{\gamma\beta}^{\delta} c_{\delta\alpha}^{\epsilon} + c_{\beta\alpha}^{\delta} c_{\delta\gamma}^{\epsilon} + c_{\alpha\gamma}^{\delta} c_{\delta\beta}^{\epsilon} \right) = 0.$$

Hence it follows that the difference of (11) and (12) is equal to zero—when we bear in mind that for arbitrary α, β, and δ it follows from

$$e_\alpha \times e_\beta = -e_\beta \times e_\alpha$$

that

$$c_{\alpha\beta}^{\delta} = -c_{\beta\alpha}^{\delta}, \tag{13}$$

and from

$$e_\alpha \times e_\alpha = 0$$

that

$$c_{\alpha\alpha}^{\delta} = 0.$$

We can now state that there exist chains of reductions that carry s_1''' and s_2''' into the same normal form s_3. The argument then continues exactly as in the preceding cases. This completes the proof of the lemma.

8. We now construct, over P, the algebra R whose basis is the set of all normal words (see § 39.4). We shall take as the *product* of the normal words w_1 and w_2 that normal form into which the word $w_1 w_2$ is carried. In virtue of the lemma, this product is uniquely defined and it is associative.

From the associative algebra R so obtained we go over to the corresponding Lie algebra $L(R)$, denoting the multiplication in the latter by the symbol $\circ$. Let us find the subalgebra of $L(R)$ generated by all the elements e_α, which, as normal words, occur in the basis of the algebras R and $L(R)$. For this purpose we show that for arbitrary α and β,

$$e_\alpha \circ e_\beta = e_\alpha \times e_\beta.$$

For if $\alpha = \beta$, then

$$e_\alpha \circ e_\alpha = 0 = e_\alpha \times e_\alpha.$$

If $\alpha > \beta$, then in virtue of (2),

$$e_\alpha \circ e_\beta = e_\alpha e_\beta - e_\beta e_\alpha = e_\beta e_\alpha + \sum_\gamma c_{\alpha\beta}^\gamma e_\gamma - e_\beta e_\alpha = \sum_\gamma c_{\alpha\beta}^\gamma e_\gamma = e_\alpha \times e_\beta.$$

If, on the other hand, $\alpha < \beta$, then in virtue of (13),

$$e_\alpha \circ e_\beta = e_\alpha e_\beta - e_\beta e_\alpha = e_\alpha e_\beta - e_\alpha e_\beta - \sum_\gamma c_{\beta\alpha}^\gamma e_\gamma = \sum_\gamma c_{\alpha\beta}^\gamma e_\gamma = e_\alpha \times e_\beta.$$

Thus, we have found in $L(R)$ a subalgebra isomorphic to L. This completes the proof of the Birkhoff-Witt Theorem. //

❖ If L is a finite-dimensional Lie algebra over a field without characteristic (see § 18.11), then there exists a finite-dimensional associative algebra R over the same field such that L can be isomorphically embedded in $L(R)$ [I. D. Ado, *Izv. Kaz. fiz. mat. ob.*, vol. 7 (1934-35), pp. 1-43; see also I. D. Ado, *Usp. Mat. Nauk*, vol. 2 (1947), pp. 159-73; Harish-Chandra, *Ann. of Math.*, vol. 50 (1949), pp. 68-76].

A theorem analogous to the Birkhoff-Witt Theorem is true for Lie rings without operators [M. Lazard, *C. R. Paris*, vol. 234 (1952), pp. 788-91]. The theorem cannot be extended to Lie rings with an arbitrary operator domain [A. I. Shirshov, *Usp. Mat. Nauk*, vol. 8 (1953), pp. 173-75]. ❖

§ 40. Derivations. Differential rings

1. When we introduced the concept of an operator ring in § 31.7, we did not use the endomorphisms of the ring R, i.e., those endomorphisms φ of the additive group of the ring that satisfies the condition

$$(ab)\varphi = a\varphi \cdot b\varphi \qquad (a, b \in R),$$

but rather we used the endomorphisms φ of the additive group that are permutable with all right and left multiplications, i.e., that satisfy the condition

$$(ab)\varphi = (a\varphi)b = a(b\varphi) \qquad (a, b \in R).$$

The choice of just these transformations can, of course, be justified on historical grounds. However, in many problems it is essential to use endomorphisms δ of the additive group of R that satisfy the condition

$$(ab)\delta = (a\delta)b + a(b\delta) \qquad (a, b \in R). \tag{1}$$

This condition is similar to the rule for finding the derivative of a product, and since δ, as endomorphism of the additive group of R, also satisfies the condition

$$(a+b)\delta = a\delta + b\delta \qquad (a, b \in R), \tag{2}$$

which is similar to the rule for finding the derivative of a sum, δ is called a *derivation* of R.

If R is an operator ring, in particular, an algebra over a commutative field, then it is natural to postulate in the definition of a derivation that δ is an operator endomorphism of the additive group. In other words, apart from (1) and (2), δ must also satisfy the condition

$$(a\alpha)\delta = (a\delta)\alpha, \tag{3}$$

where $a \in R$, and α is an arbitrary operator.

The null endomorphism of every ring R (see § 19.7) is obviously a derivation of R. An example of a non-trivial derivation is the usual differentiation in the ring of polynomials $P[x]$ in one indeterminate over a commutative field P.

2. If G is an arbitrary abelian group, then its ring of endomorphisms (see § 19.8) is associative. By replacing the operation of multiplication in this ring by the operation

$$\varphi \circ \psi = \varphi\psi - \psi\varphi,$$

we obtain, in virtue of § 8.3, a Lie ring, which we call the *Lie ring of endomorphisms* of G.

The set of derivations of an arbitrary ring R is a subring of the Lie ring of endomorphisms of the additive group of R.

For if δ_1 and δ_2 are derivations of R, then $\delta_1 + \delta_2$ is also a derivation, because for arbitrary $a, b \in R$,

$$(ab)(\delta_1+\delta_2) = (ab)\delta_1 + (ab)\delta_2 = (a\delta_1)b + a(b\delta_1) + (a\delta_2)b + a(b\delta_2) =$$
$$= (a\delta_1 + a\delta_2)b + a(b\delta_1 + b\delta_2) = [a(\delta_1+\delta_2)]b + a[b(\delta_1+\delta_2)].$$

As we have mentioned above, the null endomorphism is a derivation. The endomorphism $-\delta$ opposite to the derivation δ is itself a derivation, because for $a, b \in R$,

$$(ab)(-\delta) = -[(ab)\delta] = -[(a\delta)b + a(b\delta)] = [a(-\delta)]b + a[b(-\delta)].$$

Finally, the Lie product

$$\delta_1 \circ \delta_2 = \delta_1\delta_2 - \delta_2\delta_1$$

of the derivations δ_1 and δ_2 is itself a derivation, because for $a, b \in R$,

$$\begin{aligned}(ab)(\delta_1 \circ \delta_2) &= (ab)(\delta_1\delta_2 - \delta_2\delta_1) = [(ab)\delta_1]\delta_2 - [(ab)\delta_2]\delta_1 = \\ &= [(a\delta_1)b + a(b\delta_1)]\delta_2 - [(a\delta_2)b + a(b\delta_2)]\delta_1 = \\ &= [(a\delta_1)b]\delta_2 + [a(b\delta_1)]\delta_2 - [(a\delta_2)b]\delta_1 - [a(b\delta_2)]\delta_1 = \\ &= [(a\delta_1)\delta_2]b + (a\delta_1)(b\delta_2) + (a\delta_2)(b\delta_1) + a[(b\delta_1)\delta_2] - \\ &\quad - [(a\delta_2)\delta_1]b - (a\delta_2)(b\delta_1) - (a\delta_1)(b\delta_2) - a[(b\delta_2)\delta_1] = \\ &= [a(\delta_1 \circ \delta_2)]b + a[b(\delta_1 \circ \delta_2)].\end{aligned}$$

This proves the theorem. //

We note that if R is an algebra over a field P then the ring of operator endomorphisms of the additive group of R is also an algebra over P: if φ is such an endomorphism and if $\alpha \in P$ and $\beta \in R$, then we have to set

$$\alpha(\varphi\alpha) = (\alpha\varphi)\alpha = (\alpha\alpha)\varphi. \tag{4}$$

The Lie ring of operator endomorphisms of the additive group of R is also an algebra over P, because, by (4),

$$\begin{aligned}\alpha[(\varphi_1 \circ \varphi_2)\alpha] &= [\alpha(\varphi_1\varphi_2 - \varphi_2\varphi_1)]\alpha = \\ &= [\alpha(\varphi_1\alpha)]\varphi_2 - (\alpha\varphi_2)(\varphi_1\alpha) = \alpha[(\varphi_1\alpha) \circ \varphi_2],\end{aligned}$$

i.e.,

$$(\varphi_1 \circ \varphi_2)\alpha = (\varphi_1\alpha) \circ \varphi_2,$$

and similarly,

$$(\varphi_1 \circ \varphi_2)\alpha = \varphi_1 \circ (\varphi_2\alpha)$$

Finally, *the Lie ring of derivations of the algebra R is a subalgebra of this Lie algebra of endomorphisms*, because if δ is a derivation and $\alpha \in P$, then for $a, b \in R$,

$$(ab)(\delta\alpha) = [(ab)\delta]\alpha = [(a\delta)b + a(b\delta)]\alpha = [a(\delta\alpha)]b + a[b(\delta\alpha)].$$

This remark carries over, of course, to the case of rings with an arbitrary system of operators.

3. Let R be an associative ring. If $r \in R$, then we define a mapping δ_r by setting, for all $a \in R$,

$$a\delta_r = ar - ra. \tag{5}$$

This mapping is a derivation, because for $a, b \in R$,

$$\begin{aligned}(a + b)\delta_r &= (a + b)r - r(a + b) = a\delta_r + b\delta_r, \\ (ab)\delta_r &= abr - rab = abr - rab + arb - arb = (a\delta_r)b + a(b\delta_r).\end{aligned}$$

It is called the *inner derivation* of R determined by r. It is clear that the associative-commutative rings are the only associative rings that do not have inner derivations other than the null endomorphism.

4. If R is a Lie ring and $r \in R$, then the right multiplication by r, i.e., the mapping δ_r defined by the equation

$$a\delta_r = ar, \tag{6}$$

where $a \in R$, is a derivation of R. For we know from § 31.6 that δ_r is an endomorphism of the additive group of R. Furthermore, if $a, b \in R$, then by using the Jacobi identity and the anticommutative law (see § 8.2) we obtain:

$$(ab)\delta_r = (ab)r = -(br)a - (ra)b = (ar)b + a(br) = (a\delta_r)b + a(b\delta_r).$$

The derivation δ_r is called the *inner derivation* of the Lie ring R determined by r. Obviously, the zero rings are the only Lie rings that do not have inner derivations other than the null endomorphism ω.

The inner derivations form an ideal in the Lie ring of all derivations of the Lie ring R, so that

$$\delta_r - \delta_s = \delta_{r-s},$$
$$\delta_r \circ \delta' = \delta_{r\delta'},$$

where $r, s \in R$, and δ' is an arbitrary derivation.

For if $a \in R$, then

$$a(\delta_r - \delta_s) = ar - as = a\delta_{r-s},$$

$$a(\delta_r \circ \delta') = a(\delta_r\delta' - \delta'\delta_r) = (ar)\delta' - (a\delta')r =$$
$$= (a\delta')r + a(r\delta') - (a\delta')r = a\delta_{r\delta'}. \text{ //}$$

The mapping

$$r \to \delta_r \qquad (r \in R)$$

is a homomorphism of the Lie ring R onto the Lie ring of its inner derivations.

For if $a, r, s \in R$, then

$$a\delta_{r+s} = a(r+s) = ar + as = a\delta_r + a\delta_s = a(\delta_r + \delta_s),$$

$$a\delta_{rs} = a(rs) = -(rs)a = (sa)r + (ar)s =$$
$$= (ar)s - (as)r = a(\delta_r\delta_s - \delta_s\delta_r) = a(\delta_r \circ \delta_s). \text{ //}$$

The kernel of this homomorphism is, of course, the set of annihilators of R (see § 31.9).

The results of this subsection are reminiscent of the results of § 19.2 on the inner automorphisms of groups.

5. If R is an associative ring and L the Lie ring corresponding to it in the sense of § 8.3, then *every derivation δ of R is also a derivation of L.* Indeed,

$$(a \circ b)\delta = (ab - ba)\delta = (a\delta)b + a(b\delta) - (b\delta)a - b(a\delta) =$$
$$= (a\delta \circ b) + (a \circ b\delta).$$

The converse does not hold in general. However, *inner derivations of R are inner derivations of L, and vice versa.* For (5) can be written in the form

$$a\delta_r = a \circ r;$$

moreover, (6) must now be written in the form

$$a\delta_r = a \circ r,$$

and therefore

$$a\delta_r = ar - ra.$$

From the results of this and the preceding subsection it follows that *the inner derivations of an associative ring R form an ideal in the Lie ring of all derivations of R.*

We remark that all the results obtained in these subsections are valid for rings with an arbitrary operator domain, and in particular, for algebras.

❖ The concept of an inner derivation can be extended from the case of associative and Lie rings to arbitrary rings in such a way that the inner derivations form an ideal in the Lie ring of all derivations [R. D. Schafer, *Bull. Amer. Math. Soc.*, vol. 55 (1949), pp. 769-76]. ❖

6. A complete analogue to the concept of an operator ring is that of a differential ring. A ring R is called a *differential ring with the system of derivations Δ* if a set Δ is given and if there is assigned to every element $\delta \in \Delta$ a certain derivation of R; it is not assumed here that distinct elements of Δ must correspond to distinct derivations. We mention, without attempting a precise formulation, that it would be possible to regard Δ as a Lie ring and to define a differential ring with Δ as a Lie ring of derivations.

If R is a differential ring with a system of derivations Δ, then we must take a *differential subring* (*ideal*) of R to mean a subring (ideal) A of R such that for all $a \in A$, $\delta \in \Delta$,

$$a\delta \in A.$$

The concepts of an *isomorphism* and a *homomorphism* of two differential rings with the same system of derivations Δ are also defined in the natural way.

7. Let us examine a differential ring R with a single derivation δ, which we write in the usual notation of Analysis: if $a \in R$, then

$$a\delta = a'.$$

The equations (1) and (2) can now be rewritten in the form

$$(ab)' = a'b + ab', \tag{7}$$

$$(a + b)' = a' + b', \tag{8}$$

and hence, in particular,

$$0' = 0 \tag{9}$$

$$(-a)' = -a'. \tag{10}$$

If R has a unit element e, then by (7) we have, for $a \in R$,

$$a' = (ae)' = a'e + ae' = a' + ae',$$

so that $ae' = 0$. However, a ring with a unit element cannot have an annihilator other than zero, and therefore

$$e' = 0. \tag{11}$$

Let R be a commutative field, let $a, b \in R$ $(a \neq 0)$, and let c be the solution of the equation $ax = b$. Then

$$b' = (ac)' = a'c + ac',$$

and hence

$$c' = a^{-1}b' - a^{-1}a'c = a^{-2}(ab' - a'b). \tag{12}$$

8. An element a of R is called a *constant* relative to the derivation in question if

$$a' = 0.$$

It follows from (7)-(11) that *the constants form a subring of R which contains the unit element, provided R has one.* Thus, in the ring of integers there are no such derivations other than the null endomorphism. The same is true of the field of rational numbers, because, by (12), *in a commutative field the constants form a subfield.*

9. Let R be an associative-commutative differential ring with one derivation. We consider the ring of polynomials

$$\overline{R} = R[x, x', x'', \ldots, x^{(n)}, \ldots]$$

over R in a countable set of indeterminants (see § 8.7). We can extend the derivation given in R to the whole ring $\overline{R}$ by setting

$$(x)' = x', \quad (x^{(n)})' = x^{(n+1)} \qquad (n = 1, 2, \ldots)$$

and by making use of (7), (8), and the fact that the elements of $\overline{R}$ have unique representations in the form of polynomials. The verification that all the postulates in the definition of a derivation are then satisfied is left to the reader.

In the differential ring $\overline{R}$ so obtained the differential subring generated by R and the element x coincides with the whole of $\overline{R}$. For this reason, the differential ring R with the derivation just introduced is called the *differential ring of polynomials over R in the single indeterminate x.*

This construction can easily be carried over to the case of several indeterminates, and also to the case of several derivations.

❖ The ring $\overline{R}$ does not contain any constants other than the constants of R. In particular, if R is considered with the zero derivation, then R is the subring of constants of $\overline{R}$. ❖

CHAPTER VI

ORDERED AND TOPOLOGICAL GROUPS AND RINGS. RINGS WITH A VALUATION

§ 41. Ordered Groups

1. As a rule, the fundamental algebraic objects we have to deal with in mathematics are not groups, rings, or fields in their pure form. We have already discussed operator groups, operator rings, and differential rings: they represent a certain approximation to those concrete algebraic systems that occur in non-algebraic investigations. In the present chapter we shall follow other paths, but in the same general direction. We begin with a discussion of ordered structures.

The additive groups of integers, of rational numbers, and of real numbers are at one and the same time groups and linearly ordered sets (see § 4.1). The additive group of real functions of a real variable x defined for all values of x can be made into a partially ordered set by putting $f \leqq g$ if and only if $f(x) \leqq g(x)$ for all values of x.

In all these groups the ordering and the group operation are connected in the following way: an inequality between two elements of the group is not affected by the addition of the same element to both sides. This leads to the following definition, in which we use the multiplicative notation:

A groupoid G is called *linearly* (or *partially*) *ordered* if a linear (or partial) ordering is given for its elements such that it follows from $a \leqq b$ that $ax \leqq bx$ and $xa \leqq xb$ for all $x \in G$.

It follows from this definition that if the elements a, b, c, d of a partially ordered (in particular, a linearly ordered) groupoid G satisfy the inequality

$$a \leqq b, \qquad c \leqq d,$$

then

$$ac \leqq bd. \tag{1}$$

Obviously, if A is a subgroupoid of G, then the partial ordering of G induces a partial ordering in A.

Partially ordered groupoids G and G' are called *isomorphic* if there exists a one-to-one mapping of G onto G' that is an isomorphism both in the sense of the algebraic operation and of the partial ordering (see §§ 10.1 and 4.3). Every groupoid can be regarded as partially ordered by taking its trivial partial order (see § 4.1).

2. We define a *monotonic transformation* of a partially ordered set M as a mapping φ of M into itself such that it always follows from $a \leqq b$ that $a\varphi \leqq b\varphi$. The product of monotonic transformations is itself monotonic, so that these transformations constitute a sub-semigroup of the symmetric semigroup on M (see § 7.8), and in fact a sub-semigroup with unit element, because the identity transformation is monotonic.

If φ, ψ are monotonic transformations of a set M, then we put $\varphi \leqq \psi$ if $x\varphi \leqq x\psi$ for all $x \in M$. In this way, the semigroup of monotonic transformations is partially ordered: if χ is an arbitrary monotonic transformation, then it follows from $x\varphi \leqq x\psi$ that $x(\varphi\chi) \leqq x(\psi\chi)$ for all $x \in M$, i.e., $\varphi\chi \leqq \psi\chi$; on the other hand, from $(x\chi)\varphi \leqq (x\chi)\psi$ for all $x \in M$ it follows that $x(\chi\varphi) \leqq x(\chi\psi)$, i.e., $\chi\varphi \leqq \chi\psi$.

The following theorem is due to V.-S. Krishnan [*Bull. Soc. Math. France*, vol. 78 (1950), pp. 235-63].

Every partially ordered semigroup G with unit element can be isomorphically embedded in the partially ordered semigroup of monotonic transformations of G as a partially ordered set.

For let a transformation φ_a be assigned to every element $a \in G$ by putting

$$x\varphi_a = xa$$

for all $x \in G$. Since it follows from $x \leqq y$ that $xa \leqq ya$, the transformation φ_a is monotonic. Furthermore, we know from § 10.6 that thanks to the existence of a unit element in G the mapping $a \to \varphi_a$ $(a \in G)$ is an isomorphic mapping of G into the symmetric semigroup on G and consequently into the semigroup of monotonic transformations of G. Finally, if $a \leqq b$, then for all x we have $xa \leqq xb$, i.e., $\varphi_a \leqq \varphi_b$; conversely, from $\varphi_a \leqq \varphi_b$ it follows, in particular, that $1 \cdot a \leqq 1 \cdot b$, i.e., $a \leqq b$. //

It follows from this theorem that *every partially ordered group G can be isomorphically embedded in the partially ordered group of monotonic substitutions of G as a partially ordered set.*

3. An element a of a partially ordered group G is called *positive* if $a \geqq 1$ (or, in additive notation, $a \geqq 0$), and *negative* if $a \leqq 1$. In a linearly ordered group, every element other than the unit element is either positive or negative.

It follows from (1) that the product of positive elements of a partially ordered group is positive, so that the positive elements form a sub-semigroup

of G, which we call the *semigroup of positive elements* of the partially ordered group G. The negative elements also form a sub-semigroup.

A partially ordered group G is completely determined by its semigroup of positive elements, because $a \leqq b$ if and only if $1 \leqq ba^{-1}$.

A sub-semigroup P of a group G is the semigroup of positive elements in a suitable partial ordering of G if and only if the following conditions are satisfied:

1) $1 \in P$;

2) *If* $a \in P$ *and* $a^{-1} \in P$, *then* $a = 1$;

3) *If* $a \in P$, $x \in G$, *then* $x^{-1}ax \in P$.

For if P is the semigroup of positive elements of a partially ordered group, then 1) $1 \leqq 1$; 2) it follows from $a^{-1} \geqq 1$ that $1 \geqq a$, and since it is given that $a \geqq 1$, we have $a = 1$; 3) it follows from $a \geqq 1$ that $x^{-1}ax \geqq x^{-1}x = 1$.

Suppose, conversely, that a sub-semigroup P of a group G has the properties 1)-3). We set $a \leqq b$ if $ba^{-1} \in P$ so that, by 3), we also have

$$a^{-1}b = a^{-1}(ba^{-1})a \in P.$$

This is a partial ordering of G:

$a \leqq a$, because, by 1), $aa^{-1} = 1 \in P$;

If $a \leqq b$ and $b \leqq a$, i.e., $ba^{-1} \in P$ and $ab^{-1} = (ba^{-1})^{-1} \in P$, then, by 2), $ba^{-1} = 1$, i.e., $b = a$;

If $a \leqq b$ and $b \leqq c$, i.e., $ba^{-1} \in P$ and $cb^{-1} \in P$, then

$$ca^{-1} = (cb^{-1})(ba^{-1}) \in P, \text{ i.e., } a \leqq c;$$

Finally, if $a \leqq b$, i.e., $ba^{-1} \in P$, then $(bx)(ax)^{-1} = ba^{-1} \in P$, i.e., $ax \leqq bx$ and, by 3), $(xb)(xa)^{-1} = x(ba^{-1})x^{-1} \in P$, i.e., $xa \leqq xb$.

In this partial ordering the elements of P are the only positive elements, because $a \geqq 1$ if and only if $a\ 1^{-1} = a \in P$. //

A sub-semigroup P of a group G determines a linear ordering of the group if and only if, in addition to 1)-3), *it also satisfies the condition*

4) *For every $a \in G$ either $a \in P$ or $a^{-1} \in P$.*

For if G is linearly ordered and the element a is not positive, then $a < 1$, so that $1 < a^{-1}$, that is, the element a^{-1} is positive. Suppose, conversely, that a sub-semigroup P satisfies the conditions 1)-4) and that $a, b \in G$. If $ba^{-1} \in P$, then $a \leqq b$ in the sense of the partial ordering defined by P. But if $ba^{-1} \notin P$, then, by 4), $(ba^{-1})^{-1} = ab^{-1} \in P$, i.e., $b \leqq a$.

❖ A partially ordered set M is called *directed* if for any two elements $a, b \in M$ there exists a $c \in M$ such that

$$a \leqq c, \qquad b \leqq c. \ .$$

We can therefore speak of a *directed group*. A partially ordered group is directed if and only if it coincides with the subgroup generated by all its positive elements. ❖

4. In a partially ordered group G it follows from $a \geqq 1$ that $a^n \geqq 1$ for all natural numbers n. Therefore, because of property 2) of the semigroup of positive elements, *every strictly positive element* a (i.e., $a > 1$) *must be of infinite order*: if $a^n = 1$ $(n > 1)$, then $a^{n-1} = a^{-1}$. Hence (for the definitions, see § 9.4):

Every linearly ordered group is torsion-free.

A periodic group does not admit any partial ordering except the trivial one.

5. If two partial orderings are given in a group G, with the semigroups of positive elements P_1 and P_2, respectively, then the second partial ordering is called a *refinement* of the first if $P_2 \supset P_1$, that is, if $a \leqq b$ in the first ordering always implies the same inequality in the second ordering.

The set of all sub-semigroups of G satisfying conditions 1)-3) of § 41.3 is partially ordered by inclusion. The union of an arbitrary chain of such sub-semigroups is itself a sub-semigroup with the properties 1)-3), so that, by the Kuratowski-Zorn Theorem, every sub-semigroup with the properties 1)-3) is contained in a maximal sub-semigroup, that is, *every partial ordering of a group G can be refined to a maximal partial ordering, that is, a partial ordering without further refinements.*

If a group G admits linear orderings, then every linear ordering of G is maximal.

We shall supplement this obvious statement by a criterion for all the maximal orderings of a group to be linear [M. Ohnishi, *Osaka Math. J.* vol. 2 (1950), pp. 161-64].

6. We begin with some remarks on sub-semigroups of a group G having the properties 1)-3). If P is such a sub-semigroup, then the set P^{-1} of all elements p^{-1}, where $p \in P$, is obviously also a sub-semigroup with the properties 1)-3). Note that

$$P \cap P^{-1} = 1, \tag{2}$$

because otherwise property 2) would be violated.

LEMMA 1: *If P and Q are sub-semigroups of a group G having the properties* 1)-3) *and if*

$$P \cap Q^{-1} = 1, \tag{3}$$

then the set PQ of all elements of G that can be represented in the form pq ($p \in P$, $q \in Q$) is itself a sub-semigroup with the properties 1)-3).

For

$$(p_1q_1)(p_2q_2) = (p_1 \cdot q_1p_2q_1^{-1})(q_1q_2) = p_3q_3, \tag{4}$$

where $p_3 \in P$, $q_3 \in Q$. This shows that PQ is a sub-semigroup.

Further, $1 = 1 \cdot 1 \in PQ$. On the other hand, if $pq = 1$, then $p = q^{-1}$, and hence, by (3), $p = q = 1$. From

$$(p_1q_1)(p_2q_2) = 1$$

it therefore follows (see (4)) that $p_3 = q_3 = 1$, and since P and Q have the property 2), that

$$q_1 = q_2 = p_1 = q_1p_2q_1^{-1} = p_2 = 1.$$

Finally, for every $x \in G$,

$$x^{-1}(pq)x = (x^{-1}px)(x^{-1}qx) = p'q' \in PQ.$$

This proves the lemma. //

The intersection of an arbitrary system of sub-semigroups of a group G having the properties 1)-3) is obviously itself a sub-semigroup with these properties. Hence it follows that if an element a is contained in at least one such sub-semigroup, then there exists a minimal sub-semigroup with the properties 1)-3) containing a; we denote this minimal sub-semigroup by P_a.

7. THEOREM: *All the maximal orderings of a group G are linear if and only if:*

i) *P_a exists for every $a \in G$;*

ii) *If $b, c \in P_a$, $a \in G$, and $b \neq 1$, $c \neq 1$, then*

$$P_b \cap P_c \neq 1.$$

If a group G admits a linear ordering with the semigroup of positive elements P, then $a \in P$ or $a \in P^{-1}$, so that the condition i) is satisfied.

Suppose now that every partial ordering of a group G can be refined to a linear one. If $b, c \in P_a$ ($b \neq 1$, $c \neq 1$) but $P_b \cap P_c = 1$, then, by Lemma 1, $P_bP_c^{-1}$ is a sub-semigroup with the properties 1)-3). The partial ordering determined by this sub-semigroup can be refined, by hypothesis, to a linear ordering determined by the sub-semigroup of positive elements P. Thus,

$$b \in P, \qquad c^{-1} \in P,$$

and therefore $c \in P^{-1}$. We arrive at a contradiction to the fact that either $a \in P$, i.e., $P_a \subseteq P$ and therefore $P_a \cap P^{-1} = 1$, or that $a \in P^{-1}$, i.e., $P_a \subseteq P^{-1}$ and therefore $P_a \cap P = 1$. This proves that the condition ii) is also satisfied.

Now let us assume that a group G has the properties i) and ii) and that an element a and a sub-semigroup P with the properties 1)-3) are chosen in G.

LEMMA 2: *If* $P \cap P_a \neq 1$, *then* $P \cap P_a^{-1} = 1$.

For let $x \in P \cap P_a$, $x \neq 1$, and at the same time $y \in P \cap P_a^{-1}$, $y \neq 1$. Then $x^{-1} \in P_{a^{-1}}$, $y \in P_a^{-1} = P_{a^{-1}}$, so that by ii),

$$P_{x^{-1}} \cap P_y \neq 1.$$

However, $x^{-1} \in P^{-1}$, $y \in P$, i.e., $P_{x^{-1}} \subseteq P^{-1}$, $P_y \subseteq P$, and therefore $P^{-1} \cap P \neq 1$, in contradiction to (2). This proves the lemma.

To conclude the proof of the theorem, we choose in G an arbitrary sub-semigroup P that determines a maximal ordering. If this ordering is not yet linear, then condition 4) of § 41.3 is not satisfied, i.e., there exists an element $a \in G$ such that $a \notin P$ and $a^{-1} \notin P$. In virtue of the condition i) the sub-semigroup P_a exists, and by Lemma 2, we may assume (replacing a by a^{-1}, if necessary) that $P \cap P_a^{-1} = 1$. Therefore, by Lemma 1, the product PP_a is a sub-semigroup with the properties 1)-3) and is strictly larger than P; however, this contradicts the choice of the sub-semigroup P. This completes the proof of the theorem. //

8. The criterion just proved leads to the following theorem of E. P. Shimbireva [*Mat. Sb.*, vol. 20 (1947), pp. 145-78].

All maximal orderings of a torsion-free abelian group are linear.

For if a is an element other than 1 of a torsion-free abelian group, then the sub-semigroup P_a consists of all the powers a^n ($n = 0, 1, 2, \ldots$). Moreover, since $P_1 = 1$ always holds, condition i) is satisfied. Condition ii) is also satisfied, because if $b = a^k$, $c = a^l$ ($k \geqq 1, l \geqq 1$), then $a^{kl} \neq 1$ and $a^{kl} \in P_b \cap P_c$. //

A special case of this theorem is the Theorem of F. W. Levi [*Rend. Palermo*, vol. 35 (1913), pp. 225-36]:

Every torsion-free abelian group can be linearly ordered.

❖ Necessary and sufficient conditions for a group to be capable of a linear ordering exist, in fact, in various forms [K. Iwasawa, *J. Math. Soc. Jap.*, vol. 1 (1948), pp. 1-9; A. I. Mal'cev, *Izv. Akad. Nauk SSSR, ser. mat.*, vol. 21 (1957), pp. 199-208]. ❖

§ 42. Ordered Rings

1. All the ordered groups mentioned in the second paragraph of § 41.1 are additive groups of rings or of commutative fields. These examples of ordered rings (fields) justify the following definition, which can be stated at once for an arbitrary group with multi-operators (see § 18.1):

A ring R is called *linearly (partially) ordered* if its additive group is linearly (partially) ordered (see § 41.1), so that, by § 41.3, we can speak of positive elements and if, in addition, the product of positive elements is positive, i.e., if it follows from $a \geqq 0$, $b \geqq 0$ that $ab \geqq 0$.

It follows from this definition that the trivial partial ordering of the additive group of an arbitrary ring is a (trivial) partial ordering of the ring itself.

The condition on the product of positive elements in our definition is obviously equivalent to the fact that if $a \leqq b$ and $c \geqq 0$, then $ac \leqq bc$ and $ca \leqq cb$. Now clearly, it follows from $a \geqq 0$ that $-a \leqq 0$: we only have to add $-a$ to both sides of the first inequality; conversely, it follows from $a \leqq 0$ that $-a \geqq 0$. Hence we have the following *rule of signs*:

$$\text{If } a \leqq 0,\ b \geqq 0, \text{ then } ab \leqq 0 \text{ and } ba \leqq 0;$$
$$\text{If } a \leqq 0,\ b \leqq 0, \text{ then } ab \geqq 0.$$

Hence it follows that it would have been impossible to define an ordered ring as a ring with an ordering in which both the additive group and the multiplicative groupoid of the ring are ordered in the sense of § 41.1.

We mention that in the case of rings without divisors of zero the condition on products of positive elements in the definition of an ordered ring can be given in the following amended form: the product of strictly positive elements is strictly positive, i.e., it follows from $a > 0$, $b > 0$ that $ab > 0$. This is equivalent to the fact that if $a < b$ and $c > 0$, then $ac < bc$ and $ca < cb$.

Finally, an *isomorphism* of ordered rings is defined essentially as was done in § 41.1 for ordered groupoids.

2. By § 41.3, every partial ordering of a ring R is determined by the additive semigroup of positive elements P: $a \leqq b$ if and only if $b - a \in P$. From § 41.3 and the definition of an ordered ring we derive the following:

A sub-semigroup P of the additive group of a ring R is the semigroup of positive elements in some partial ordering of the ring if and only if the following conditions are satisfied:

1′) $0 \in P$;

2′) *If* $a \in P$ *and* $-a \in P$, *then* $a = 0$;

3′) *If* $a \in P$ *and* $b \in P$, *then* $ab \in P$.

A partial ordering of a ring R determined by an additive semigroup P with the properties 1')-3') *is linear if and only if the following additional condition holds*:

4') *For every* $a \in R$ *either* $a \in P$ *or* $-a \in P$. .

3. Let $a_1, a_2, \ldots, a_n$ be non-zero elements of a ring R. Henceforth their *product* will be taken to mean every product (in the non-associative case, with a suitable distribution of parentheses) in which each factor is one of the a_i $(i = 1, 2, \ldots, n)$ and where each of the a_i may occur in the product several times or not at all. A product of elements $a_1, a_2, \ldots, a_n$ is called *even* if every a_i $(i = 1, 2, \ldots, n)$ occurs in it an even number of times. Finally, in speaking of a *sum of products* of certain elements we shall assume that all these products occur in the sum in question with the plus sign.

The following theorem is due to R. E. Johnson [*Proc. Amer. Math. Soc.*, vol. 3 (1952), pp. 414-16; see also V. D. Podderjugin, *Usp. Mat. Nauk*, vol. 9 (1954), pp. 211-16].

A ring R is a ring without divisors of zero and admits of linear ordering if and only if every sum of even products of its elements is different from zero.

For suppose that R is a linearly ordered ring without divisors of zero and that an even product of its elements is given. This product does not change if all its negative factors are replaced by their opposite elements, so that it is equal to a product of strictly positive elements, and hence, in virtue of the absence of divisors of zero, it is itself strictly positive. Consequently, every sum of even products is strictly positive and therefore different from zero.

4. Now let R be a ring in which every sum of even products of its elements is different from zero. Then R cannot have divisors of zero, because it would follow from $ab = 0$, $a \neq 0$, $b \neq 0$ that the even product $(ab)(ab)$ is zero.

We shall call a subset Q of our ring R a *regular* subset if it satisfies the following conditions:

1) $0 \notin Q$;

2) If $a \in Q$, then $-a \notin Q$;

3) If $a_1, a_2, \ldots, a_n \in Q$ $(n \geqq 0)$ and $x_1, x_2, \ldots, x_k$ are non-zero elements of R $(k \geqq 0)$ with $n + k > 0$ and if σ is a sum of products of all these elements, then the x_j can be replaced in at least one way by elements x_j', where

$$x_j' = x_j \quad \text{or} \quad x_j' = -x_j \qquad (j = 1, 2, \ldots, k), \tag{1}$$

in such a way that after the replacement the sum σ is different from zero.

LEMMA 1: *The empty set is regular.*

For suppose that a sum σ of products of non-zero elements $x_1, x_2, \ldots, x_k \in R$ $(k \geqq 1)$ remains equal to zero under an arbitrary replacement of the form (1). If the element x_1 occurs in each term of this sum an odd number of times, then σx_1 is a sum in which each term contains the factor x_1 an even number of times, and this sum again remains equal to zero no matter what replacement of the form (1) is made.

On the other hand, suppose that some terms of the sum contain the factor x_1 an even number of times and other terms, an odd number of times. We denote the sum of the first set of terms by σ_1 and that of the second, by σ_2:

$$\sigma = \sigma_1 + \sigma_2$$

Thus, the sum $\sigma_1 + \sigma_2$ is zero under an arbitrary replacement of the form (1). Consequently the same is true of the sum obtained from it by replacing x_1 by $-x_1$; this sum can obviously be written in the form $\sigma_1 - \sigma_2$; and hence the same is true also of the sum $2\sigma_1$, in which each term contains the factor x_1 an even number of times.

When we perform these transformations in succession for each of the elements $x_1, x_2, \ldots, x_k$, we eventually arrive at a sum of even products of these elements that is equal to zero. But this contradicts our assumption, and our lemma is proved.

LEMMA 2: *If a regular set Q contains neither a nor $-a$, where $a \neq 0$, then at least one of the unions $Q \cup a$, $Q \cup (-a)$ is regular.*

For suppose that σ_1 is a sum of products of elements of Q, of a, and of non-zero elements

$$x_1, x_2, \ldots, x_k, \tag{2}$$

such that σ_1 is zero no matter what replacement of the form (1) is performed on the elements (2). Furthermore, let σ_2 be a sum of products of elements of Q, of a, and of non-zero elements

$$y_1, y_2, \ldots, y_l, \tag{3}$$

that remains equal to zero when a is replaced by $-a$ and the elements (3) are subject to any replacement of the form (1). Then $\sigma_1\sigma_2$ is a sum of products of elements of Q and of the elements

$$a, x_1, x_2, \ldots, x_k, y_1, y_2, \ldots, y_l, \tag{4}$$

that is equal to zero under every replacement of the form (1) applied to the elements (4). However, this contradicts the regularity of Q, and the lemma is proved.

5. The union of every chain of regular subsets of our ring R is obviously regular. Therefore, by the Kuratowski-Zorn Theorem, R has maximal regular subsets. Let Q be one of them. By Lemma 2, for every $a \in R$ $(a \neq 0)$ either a or $-a$ belongs to Q.

If $a, b \in Q$, then $a + b \in Q$. For $a + b \neq 0$ in virtue of 2) in the definition of a regular set. If the element $c = -(a + b)$ were contained in Q, then we would have the equation

$$a + b + c = 0,$$

in contradiction to the regularity of Q.

If $a, b \in Q$, then $ab \in Q$. For $ab \neq 0$ in virtue of 1) in the definition of a regular set and because of the absence of divisors of zero. If the element $d = -ab$ were contained in Q, then we would have the equation

$$ab + d = 0,$$

in contradiction to the regularity of Q.

Hence it follows that the union $P = Q \cup 0$ is an additive sub-semigroup of R satisfying all the conditions 1')-4') of § 42.2. This completes the proof of the theorem of § 42.3. //

6. From this theorem we derive the following:

A commutative-associative ring R is an integral domain admitting a linear ordering if and only if no sum of squares of non-zero elements of R can be equal to zero.

Hence it follows that the field of complex numbers cannot have a linear ordering: in this field we have the equation

$$1^2 + i^2 = 0.$$

❖ The field of power series over an ordered field P (see § 11.7), and hence also the field of rational fractions over P, can be linearly ordered. ❖

7. *If a linearly ordered ring R has a unit element, then it contains a subring isomorphic to the ring of integers with its usual ordering.*

For the unit element must be strictly positive, because it coincides with its square. Therefore all the positive multiples of the unit element are strictly positive, and by § 41.4, they are all distinct.

Hence it follows that *the usual ordering of the ring of integers is its only possible linear ordering.*

❖ The ring of integers is the only linearly ordered ring without divisors of zero and with a unit element in which the ordered subset of positive elements is well-ordered (see § 5.4). ❖

8. *An ordering of an integral domain R can be extended in a unique way to its field of fractions $\bar{R}$.*

Let us assume first of all that the ordering of the ring R is already extended to the field $\bar{R}$. It follows from $b \neq 0$ that $b^2 > 0$, and hence, if $a/b > 0$, we find that

$$(a/b) \cdot b^2 = ab > 0.$$

Conversely, if $ab > 0$ in R, then since $(b^{-1})^2 > 0$, we have

$$ab \cdot (b^{-1})^2 = a/b > 0.$$

This shows the uniqueness of a possible extension of the ordering from R to $\bar{R}$.

To prove that such an extension exists, we set $a/b > 0$ if and only if $ab > 0$. If $a/b = c/d$, i.e., $ad = bc$ (see § 11.2), then $abd^2 = b^2cd$. Therefore, since $d^2 > 0$, $b^2 > 0$, it follows from $ab > 0$ that $cd > 0$, i.e., that $c/d > 0$; consequently, our definition of strictly positive elements in $\bar{R}$ is legitimate.

If $a, b \in R$ and $a \neq 0$, $b \neq 0$, then only one of the two elements ab and $(-a)b$ is strictly positive in R, and hence either $a/b > 0$ or $-(a/b) = (-a)/b > 0$ (see § 11.4), and these cases are mutually exclusive.

If $a/b > 0$, $c/d > 0$, i.e., $ab > 0$, $cd > 0$, then $abd^2 > 0$, $b^2cd > 0$, and hence

$$(ad + bc)bd = abd^2 + b^2cd > 0,$$

so that (see § 11.4)

$$a/b + c/d > 0.$$

Under the same assumption, we have

$$(ac)(bd) = (ab)(cd) > 0,$$

so that (see § 11.2)

$$(a/b) \cdot (c/d) > 0.$$

This shows that the set P consisting of the zero element and all the strictly positive elements of $\bar{R}$ satisfies all the conditions of § 42.2, so that it determines a linear ordering in $\bar{R}$. This ordering is an extension of the given linear ordering of R, because by § 11.2,

$$a = (ab)/b \qquad (b \neq 0),$$

but $a > 0$ implies that $(ab)b = ab^2 > 0$, and vice versa. //

From this theorem and § 42.7 it follows that *the usual ordering of the field of rational numbers is its only possible linear ordering.*

Since the additive group of an arbitrary linearly ordered field K is, by § 41.4, torsion-free, K is a field without characteristic (see § 18.11), so that in virtue of what has been shown above *every linearly ordered field K contains the field of rational numbers in its usual ordering as its prime subfield.*

9. *The ordered set Q of strictly positive elements of an arbitrary linearly ordered associative field K constitutes a linearly ordered group under multiplication.*

For, from § 42.7, we know that the unit element of K belongs to Q. Furthermore, if $a \in Q$, i.e., $a > 0$, then it follows that $a^{-1} \in Q$, because $a^{-1} < 0$ would imply, by § 42.1, that $aa^{-1} = 1 < 0$. The set Q therefore proves to be a multiplicative group. The remaining statements of the theorem follow from the definition of an ordering of a ring without divisors of zero (see § 42.1). //

❖ Every linearly ordered group can be isomorphically embedded in the linearly ordered multiplicative group of strictly positive elements of a linearly ordered associative field [A. I. Mal'cev, *Dokl. Akad. Nauk*, vol. 60 (1948), pp. 1499-1501; B. H. Neumann, *Trans. Amer. Math. Soc.*, vol. 66 (1949), pp. 202-52]. ❖

§ 43. Archimedean Groups and Rings

1. A subset N of a partially ordered set M is called *convex* if, whenever it contains any two comparable elements a and b $(a < b)$, it also contains all the elements x of M that satisfy the inequalities $a \leqq x \leqq b$.

A subgroup A of a partially ordered group G is convex if and only if, whenever it contains a positive element a, it also contains all the positive elements x of G that satisfy the inequality $x \leqq a$.

For, the inequalities $a \leqq x \leqq b$ and $1 \leqq a^{-1}x \leqq a^{-1}b$ are equivalent. //

Let G and G' be partially ordered groups; let P and P' be their semigroups of positive elements; and let φ be a mapping of G onto G'. We call φ a *monotonic homomorphism* if it is a homomorphism in the group-theoretical sense and if, furthermore, $P\varphi = P'$. It is not asserted, of course, that P is the complete inverse image of P': we recall 2) of § 41.3.

The kernel A of a monotonic homomorphism φ is a convex normal subgroup.

For suppose that $a \in A$, $a \geqq 1$, and $x \in G$, $1 \leqq x \leqq a$. From $x \geqq 1$ and the definition of φ it follows that

$$x\varphi \geqq 1'. \tag{1}$$

On the other hand, since $x^{-1}a \geqq 1$ and $a\varphi = 1'$, we have

$$1' \leqq (x^{-1}a)\varphi = (x\varphi)^{-1}(a\varphi) = (x\varphi)^{-1}. \tag{2}$$

From (1) and (2) it follows that $x\varphi = 1'$, i.e., that $x \in A$.

If A is a convex normal subgroup of a partially ordered group G, then the factor group $G' = G/A$ can be partially ordered in such a way that the natural mapping of G onto G' is a monotonic homomorphism.

For if P is the semigroup of positive elements of G, we denote by P' the collection of cosets of A that contain at least one element of P. The theorem will be proved if we can show that P' is a sub-semigroup of G' and has the properties 1)-3) of § 41.3. In fact, we need only verify property 2), because the remaining statements are obvious.

Suppose that $p \in P$, or $pA \in P'$, and that, in addition, $(pA)^{-1} = p^{-1}A \in P'$. An element $p_0 = p^{-1}a$ therefore exists such that $p_0 \in P$, $a \in A$. Hence $pp_0 = a$, that is, $a \geqq 1$, and since $1 \leqq p_0$, we have

$$p \leqq pp_0 = a.$$

In view of the inequality $1 \leqq p$ and the fact that A is convex, it follows from this that $p \in A$, i.e., $pA = A$, as required. //

2. The convex subgroups of an arbitrary partially ordered group G include G itself and the unit subgroup E. The definition of a convex subgroup shows that the intersection of every set of convex subgroups and the union of an arbitrary chain of convex subgroups are themselves convex.

The convex subgroups of a linearly ordered group G form a chain under set-theoretical inclusion.

For let A and B be convex subgroups of a group G, where B contains an element b not in A. Without loss of generality, we may assume that $b > 1$. If A were to contain an element a such that $b < a$, then by the convexity of A we would have $b \in A$, contrary to assumption. Thus, all the positive elements of A are less than b, and since B is convex they all belong to B, so that $A \subset B$. //

If a is a strictly positive element of a linearly ordered group G, then the set A consisting of all the elements x such that $1 \leqq x \leqq a^n$ for some natural number n and the inverses of these elements is the minimal convex subgroup containing a.

All the elements in question must, in fact, belong to every convex subgroup containing a. Suppose, further, that

$$\left.\begin{aligned} 1 &\leqq x \leqq a^k, \\ 1 &\leqq y \leqq a^l \end{aligned}\right\} \tag{3}$$

for certain natural numbers k and l. Then by (1) of § 41.1,

$$1 \leqq xy \leqq a^{k+l},$$

so that $xy \in A$ and also $(xy)^{-1} = y^{-1}x^{-1} \in A$. On the other hand, it follows from (3) that

$$a^{-l} \leqq y^{-1} \leqq 1,$$

and therefore

$$a^{-l} \leqq xy^{-1} \leqq a^{k}.$$

Thus, if $1 \leqq xy^{-1}$, then $xy^{-1} \in A$. But if $xy^{-1} < 1$, then

$$1 < (xy^{-1})^{-1} \leqq a^{l},$$

that is, $(xy^{-1})^{-1} \in A$, so that we also have $xy^{-1} \in A$. It is now obvious that A is a subgroup, contains a, and is convex. //

3. Two strictly positive elements a and b of a linearly ordered group G are said to belong to the same *archimedean class* if they generate one and the same convex subgroup of G. All the strictly positive elements, therefore, split into disjoint archimedean classes. It is convenient to assume that the element 1 forms an archimedean class by itself.

The set of archimedean classes of G can be linearly ordered in accordance with the natural linear order of jumps (that is, of pairs of adjacent subgroups) in the chain of convex subgroups of G.

A linearly ordered group G is called *archimedean* if it has no convex subgroups other than G and E. From the last theorem of § 43.1 we derive the following:

A linearly ordered group G is archimedean if and only if for every pair a, b of strictly positive elements we can find a natural number $n = n(a, b)$ such that $a^n > b$.

The additive group of the real numbers in their natural order and the subgroups of this group are examples of archimedean groups.

4. HÖLDER'S THEOREM: *Every archimedean group is commutative and is isomorphic to a subgroup of the additive group of real numbers in its natural ordering.*

Suppose that an archimedean group G has a minimal strictly positive element a. If b is an arbitrary strictly positive element of G, then there exists a natural number n such that

$$a^n \leqq b < a^{n+1}.$$

Hence

$$1 \leqq ba^{-n} < a,$$

and therefore $ba^{-n} = 1$, so that $b = a^n$. All the strictly positive elements of G are therefore expressible as powers of a, and since by § 41.1,

$$1 < a < a^2 < \ldots < a^n < \ldots,$$

G proves to be isomorphic to the additive group of integers in its natural ordering.

Let us now assume that there is no minimal element among the strictly positive elements of an archimedean group G. In that case, for every $a > 1$ there exists a $b > 1$ such that $b^2 \leqq a$. For there exists an element c such that

$$1 < c < a,$$

and hence $1 < c^{-1}a$. If $c^2 \leqq a$, then $b = c$. If, on the other hand, $c^2 > a$, then $1 > c^{-1}ac^{-1}$, so that $a > (c^{-1}a)^2$, and therefore $b = c^{-1}a$.

To prove that G is commutative, it is sufficient to show that any two of its strictly positive elements a, b are permutable. If $ab \neq ba$, then without loss of generality we may assume that

$$c = a^{-1}b^{-1}ab > 1.$$

As we have shown, there exists an element d such that $1 < d$ and $d^2 \leqq c$. We may assume here that $d < a$ and $d < b$, so that there exist natural numbers k and l for which

$$d^k \leqq a < d^{k+1}, \qquad d^l \leqq b < d^{l+1}.$$

Hence

$$ab < d^{k+l+2},$$

$$a^{-1}b^{-1} \leqq d^{-(k+l)},$$

so that

$$c < d^2,$$

and since by assumption $d^2 \leqq c$, we have arrived at a contradiction. This proves that G is commutative.

5. We find it convenient now to go over to the additive notation of the group operation in G. We fix a strictly positive element e in G and assign to every element $a \in G$ the class U_a of all those rational numbers m/n $(n > 0)$ for which

$$me \leqq na. \tag{4}$$

If $m/n \in U_a$ and $p/q \leqq m/n$, then we also have $p/q \in U_a$, because it follows from (4) and $pn \leqq qm$ that

$$(pn)e \leqq (qm)e \leqq (qn)a,$$

and hence

$$pe \leqq qa.$$

In particular, if $m/n \in U_a$ and $p/q = m/n$, then we also have $p/q \in U_a$, that is, the class U_a can, in fact, be regarded as a set of rational numbers.

U_a does not contain all the rational numbers: since G is archimedean, there exists a natural number m such that $me > a$, and hence $m \in U_a$. Consequently, U_a is the lower class of a cut in the system of rational numbers, and so, by the Dedekind theory, it determines a real number a. We set

$$a = a\theta.$$

6. If $a \leqq b$, then $a\theta \leqq b\theta$, because from (4) and $a \leqq b$ it follows that

$$me \leqq nb,$$

i.e., $U_a \subseteq U_b$.

Next, we shall show that

$$(a + b)\theta = a\theta + b\theta. \tag{5}$$

In fact, if

$$m/n \in U_a, \qquad p/q \in U_b,$$

i.e.,

$$me \leqq na, \qquad pe \leqq qb,$$

then

$$(mq + np)e \leqq nq(a + b),$$

and hence

$$(mq + np)/nq = m/n + p/q \in U_{a+b}.$$

On the other hand, it can be shown in the same way that if

$$m/n \in U_a, \qquad p/q \in U_b,$$

then

$$m/n + p/q \in U_{a+b},$$

and so (5) follows.

The mapping θ is therefore a monotonic homomorphism of G onto a subgroup of the additive group of real numbers. In this homomorphism only the zero element of G goes over into zero: if $a > 0$, then since G is archimedean, there exists a natural number m such that $ma > e$, that is, $1/m \in U_a$, and therefore $a\theta > 0$. The homomorphism θ is consequently an isomorphism and it preserves the relation of order. //

Note that the subgroup $G\theta$ of the additive group of real numbers onto which G has been mapped isomorphically contains the number 1, because $e\theta = 1$.

7. LEMMA: *If A and B are subgroups of the additive group of real numbers in its natural ordering and if φ is a monotonic homomorphism of A onto B, then there exists a real number r $(r \geqq 0)$ such that for all $a \in A$,*

$$a\varphi = ar. \tag{6}$$

For if there exists an $a \in A$ $(a > 0)$ such that $a\varphi = 0$, then we also have $(na)\varphi = 0$ for all natural numbers n, so that $B = A\varphi = 0$. In that case, $r = 0$.

Let us therefore assume that φ is an isomorphism. If $a_1, a_2 \in A$, $a_1 > 0$, $a_2 > 0$, then $a_1\varphi > 0$, $a_2\varphi > 0$. If $a_1\varphi / a_2\varphi < a_1/a_2$, then there exists a positive rational number m/n such that

$$a_1\varphi / a_2\varphi < m/n < a_1/a_2. \tag{7}$$

Hence $na_1 > ma_2$, or $n(a_1\varphi) > m(a_2\varphi)$, although it follows from (7) that $n(a_1\varphi) < m(a_2\varphi)$. We arrive at the same contradiction when we assume that $a_1\varphi / a_2\varphi > a_1/a_2$. Thus,

$$a_1\varphi / a_2\varphi = a_1/a_2,$$

and then

$$r = a\varphi / a \qquad (a \in A;\ a > 0).$$

In fact, if $a' < 0$, then

$$a'\varphi = -(-a')\varphi = -(-a')r = a'r,$$

so that the number r satisfies the condition (6) for all elements of A. //

We call a homomorphism φ of a linearly ordered group G onto a linearly ordered group G' an *inversion homomorphism* if it follows from $a \in G$ $(a \geqq 0)$ that $a\varphi \leqq 0$. *The lemma just proved is also valid if φ is an inversion homomorphism; in that case, however, $r \leqq 0$.*

For if φ is an inversion homomorphism of A onto B, then the mapping $-\varphi$, where

$$a(-\varphi) = -a\varphi \qquad (a \in A),$$

is a monotonic homomorphism. Consequently, there exists a real number r $(r \geqq 0)$ such that

$$a(-\varphi) = ar \qquad (a \in A);$$

but then

$$a\varphi = -[a(-\varphi)] = -ar = a(-r) \qquad (a \in A),$$

where $-r \leqq 0$.

8. A linearly ordered ring R is called *archimedean* if its additive group is archimedean. The following theorem is due to Ja. V. Hion [*Usp. Mat. Nauk*, vol. 9 (1954), pp. 237-42; see also G. T. Tallini, *Atti Accad. Naz. Lincei, Rend.*, vol. 18 (1955), pp. 367-73].

Every archimedean ring R is associative and commutative. Furthermore, it is either a zero ring on a subgroup of the additive group of real numbers or it is isomorphic to a subring of the field of real numbers in its natural ordering.

For by virtue of Hölder's Theorem (see § 43.4), we may assume that the additive group of R is a subgroup of the additive group of real numbers. However, the multiplication in R need not, in general, coincide with the usual multiplication of real numbers, and in contrast to the latter, we shall denote it by $a \cdot b$.

If $a \in R$, then the transformation $x \to x \cdot a$ $(x \in R)$ is a monotonic endomorphism of the additive group of R when $a \geqq 0$, and an inversion endomorphism when $a \leqq 0$. Therefore, by § 43.7, there exists a real number r_a such that

$$x \cdot a = x r_a \qquad (x \in R),$$

where $a \geqq 0$ implies that $r_a \geqq 0$.

Since for all $x \in R$,

$$x \cdot (a + b) = x \cdot a + x \cdot b = x r_a + x r_b = x(r_a + r_b),$$

i.e., $r_{a+b} = r_a + r_b$, the mapping $a \to r_a$ is a monotonic homomorphism of the additive group of R into the additive group of real numbers. Therefore, by the lemma, there exists a real number s $(s \geqq 0)$ such that for all $a \in R$,

$$r_a = as.$$

If $s = 0$, then $r_a = 0$ for all a, so that R is a zero ring. But if $s > 0$, then the mapping $a \to r_a$ is a monotonic isomorphism of the additive group. Also, it follows from

$$a \cdot b = a r_b = a(bs) = (ab)s$$

that

$$r_{a \cdot b} = (a \cdot b)s = [(ab)s]s = (as)(bs) = r_a r_b.$$

Thus, for $s > 0$ the correspondence $a \to r_a$ maps the ring R monotonically and isomorphically onto a subring of the field of real numbers in its natural ordering. //

§ 44. Rings With a Valuation

1. In the field of real numbers and in the field of complex numbers we have the concept of the absolute value, or modulus. Recalling the usual properties that attach to this concept and bearing in mind, in particular, that the absolute value is a non-negative real number and that the field of real numbers is linearly ordered, we arrive at the following general definition:

Let W be a linearly ordered ring (see § 42.1). A ring R is said to have a *valuation, with values in* W, if to every element $a \in R$ there is assigned an element $w(a) \in W$, the *value* of a, and if the following conditions are satisfied:

1) $w(0) = 0$; $w(a) > 0$ for $a \neq 0$;

2) $w(ab) = w(a)w(b)$;

3) $w(a - b) \leqq w(a) + w(b)$.

It follows from 3) and 1) that for all $b \in R$ we have $w(-b) \leqq w(b)$, and since $b = -(-b)$, we have

3_1) $w(-b) = w(b)$ $(b \in R)$.

Hence it follows that

3_2) $w(a + b) \leqq w(a) + w(b)$.

Examples of fields with a real valuation, apart from the field of real numbers and that of complex numbers, are the quaternions and the Cayley algebra. For in these fields we can take the value of an element a to be the square root of its norm $n(a)$, as introduced in §§ 36.9 and 36.11, respectively. The norm so defined is the length of the vector a in a four-dimensional and an eight-dimensional euclidean space, respectively, and therefore condition 3) is satisfied. Condition 1) is obvious; and the validity of condition 2) follows from the equations (8) and (15) of § 36.

Every linearly ordered ring R can be given a valuation, with values in R itself.

For we define the *absolute value* $|a|$ of an element $a \in R$ as the positive one of the two elements a and $-a$. The absolute value satisfies the conditions in the definition of a valuation. For the condition 1) is obvious. Condition 2) easily follows from the rule of signs (see § 42.1):

If $a \geqq 0$, $b \geqq 0$, then $|a| \cdot |b| = ab = |ab|$;
If $a \leqq 0$, $b \geqq 0$, then $|a| \cdot |b| = (-a)b = -ab = |ab|$;
If $a \leqq 0$, $b \leqq 0$, then $|a| \cdot |b| = (-a)(-b) = ab = |ab|$.

Let us show that condition 3) is also satisfied. If $a \geqq b \geqq 0$, then $b \geqq -b$ and $a + b \geqq a - b$, and therefore

$$|a-b| = a-b \leqq a+b = |a| + |b|.$$

If $a \geqq 0 \geqq b$, then

$$|a-b| = a-b = a+(-b) = |a| + |b|.$$

If $0 \geqq a \geqq b$, then $-a \geqq a$ and $-a-b \geqq a-b$, and therefore

$$|a-b| = a-b \leqq -a-b = |a| + |b|.$$

If $b \geqq a \geqq 0$, then $a \geqq -a$ and $a+b \geqq b-a$, and therefore

$$|a-b| = -(a-b) = b-a \leqq a+b = |a| + |b|.$$

If $b \geqq 0 \geqq a$, then

$$|a-b| = b-a = b+(-a) = |a| + |b|.$$

Finally, if $0 \geqq b \geqq a$, then $b \leqq -b$ and $b-a \leqq -a-b$, and therefore

$$|a-b| = b-a \leqq -a-b = |a| + |b|. \text{ //}$$

2. In all the examples discussed above, all the positive elements of W are values of elements of R. In such a case we shall say that W is a *valuation ring*.

Let R be a ring with a valuation and with W as valuation ring. *If R has a unit element* 1, *then $w(1)$ is the unit element of W.* For if $a \in R$, then

$$w(a) = w(a \cdot 1) = w(a) \cdot w(1),$$

so that $w(1)$ is a unit element for all positive elements of W and hence also for the elements opposite to them.

The rings R and W either both have, or both do not have divisors of zero.

For if $a, b \in R$, $a \neq 0$, $b \neq 0$, but $ab = 0$, then $w(a) > 0$, $w(b) > 0$, but

$$w(a)w(b) = w(ab) = 0.$$

Conversely, if W has divisors of zero, then we can find strictly positive elements α and β in W such that $\alpha\beta = 0$. Then R contains elements a and b such that $w(a) = \alpha$, $w(b) = \beta$, and hence $a \neq 0$, $b \neq 0$. However,

$$w(ab) = w(a)w(b) = \alpha\beta = 0,$$

and hence $ab = 0$.

If R is a commutative or an associative ring, then the same is true of W.

For the validity of the commutative or the associative law for positive elements of W follows immediately from the condition 2) in the definition

of a ring with a valuation, and all other elements of W are opposites of positive elements.

From these arguments we derive the following statement:

If R is a division ring or a ring with unique division or a field, then the same is true of W.

3. If a linearly ordered ring W has a unit element 1, then every ring R admits the following *trivial valuation* with values in W:

$$w(0)=0; \quad w(a)=1 \text{ for all } a\in R \ (a\neq 0).$$

It is clear that all the conditions 1)-3) of § 44.1 are satisfied. In fact, even the following condition, which is stronger than 3), holds:

3′) $w(a-b)\leqq \max(w(a), w(b))$.

Hence it follows, as before, that

$3_1'$) $w(-b)=w(b)$,

and therefore

$$w(a+b)=w[a-(-b)]\leqq \max(w(a), w(-b))=\max(w(a), w(b)),$$

or

$3_2'$) $w(a+b)\leqq \max(w(a), w(b))$.

Valuations subject to the conditions 1), 2), and 3′) are called *non-archimedean* in contrast to an *archimedean valuation*, in which condition 3′) does not hold and only condition 3) holds. The usual valuations of the field of real numbers, the field of complex numbers, the field of quaternions, and the Cayley algebra (see § 44.1) are archimedean. So is the valuation of a linearly ordered ring described in § 44.1.

4. In the definition of a non-archimedean valuation—that is, in the axioms 1), 2), and 3′)—we have only made use of the ordering and of the multiplication of positive elements of W. This definition can therefore be carried over word for word to the case of a valuation of a ring R by the elements of a linearly ordered multiplicative groupoid G with a zero—that is, with an element 0 such that for all $x\in G$, $0\leqq x$, we have

$$x\cdot 0=0\cdot x=0.$$

This is precisely the meaning that we shall attach henceforth to a *non-archimedean valuation* of a ring R. In that case, we shall speak of a valuation of R with the *value groupoid* G, provided that every element of G is the value of some element of R.

Let $w(a)$ be a non-archimedean valuation of a ring R with G as value groupoid. In that case we have results similar to those of § 44.2, and they can be proved by the same arguments. We may also note that in the case in question the valuation determines a homomorphic mapping of the multiplicative groupoid of R onto G.

If R has the unit element 1, *then $w(1)$ is the unit element of G.*

The ring R and the groupoid G either both have, or both do not have divisors of zero: the meaning of a divisor of zero in a groupoid with a zero element is obvious.

If R is an associative ring, then G is a semigroup. If R is commutative, then so is G.

If R is an associative field, then the non-zero elements of G form a group under the multiplication given in G.

5. Let R be a linearly ordered ring and G, the ordered set of archimedean classes of the additive group (see § 43.3) of R. We shall denote the class containing the element $a \in R$, $a \geqq 0$, by $\bar{a}$. If $a, b \in R$, $a \geqq 0$, $b \geqq 0$, so that $ab \geqq 0$, then we set

$$\bar{a}\bar{b} = \overline{ab}. \tag{1}$$

The equation (1) *turns the set G into a linearly ordered groupoid with zero, the* so-called *groupoid of archimedean classes* of R.

For if $\bar{a}_1 = \bar{a}_2$, then, say, $a_1 \leqq a_2$. However, there exists a natural number n such that $a_2 \leqq na_1$, and therefore, because $b \geqq 0$,

$$a_1 b \leqq a_2 b \leqq n(a_1 b),$$

and hence $\overline{a_1 b} = \overline{a_2 b}$. This shows that the equation (1) is indeed a definition of multiplication in G.

Further, if $\bar{a}_1 < \bar{a}_2$, then $a_1 < a_2$, and so for every $b \geqq 0$,

$$a_1 b \leqq a_2 b, \qquad b a_1 \leqq b a_2,$$

and hence for every $\bar{b} \in G$,

$$\bar{a}_1 \bar{b} \leqq \bar{a}_2 \bar{b}, \qquad \bar{b}\bar{a}_1 \leqq \bar{b}\bar{a}_2.$$

Thus, the groupoid G proves to be linearly ordered (see § 41.1).

Finally, the class $\bar{0}$ is obviously the zero of the linearly ordered groupoid G. //

The following theorem is due to Ja. V. Hion [*Izv. Akad. Nauk SSSR, ser. mat.*, vol. 21 (1957), pp. 311-28].

Every linearly ordered ring R admits a non-archimedean valuation, with its groupoid of archimedean classes as value groupoid.

For every $a \in R$ we choose as $w(a)$ the archimedean class that contains the absolute value $|a|$ (see § 44.1). Condition 1) obviously holds, and the validity of condition 2) follows from (1) and from the fact that this condition holds for the absolute values. Let us show that condition 3′) also holds. Indeed, since $|-b| = |b|$, we have by § 44.1 that

$$|a-b| = |a+(-b)| \leqq |a| + |b| \leqq 2\max(|a|, |b|),$$

and therefore

$$|\overline{a-b}| \leqq \max(|\bar{a}|, |\bar{b}|),$$

i.e.,

$$w(a-b) \geqq \max(w(a), w(b)).$$

6. Let R be a ring with a non-archimedean valuation in a linearly ordered groupoid G with a zero. We shall use the multiplicative notation in G, whereas up to now we have preferred the additive notation; furthermore, we replace the ordering in G by its inverse (see § 4.6). Naturally, the element 0 will now be denoted by the symbol ∞; and for all $x \in G$, we have

$$\infty \geqq x, \quad x + \infty = \infty + x = \infty. \tag{2}$$

In this way, the non-archimedean valuation given in R is converted into a so-called *logarithmic valuation*, with values in the additive groupoid G. The conditions 1), 2), and 3′) now take the form:

1_0) $w(0) = \infty$; $w(a) < \infty$ for $a \neq 0$;

2_0) $w(ab) = w(a) + w(b)$;

$3_0'$) $w(a-b) \geqq \min(w(a), w(b))$.

The converse process, of course, is also possible.

If a non-archimedean valuation with (non-negative) real values is given in R, then the logarithmic valuation obtained from it can also be taken to be real—specifically, with values in the ordered additive groupoid of real numbers, supplemented with the symbol ∞. For, the transformation of the groupoid of values described above can be obtained in our case by replacing every positive real number α by $-\log \alpha$.

Let us choose a commutative field P, and let us take for R any of the following three rings: *the ring of polynomials* $P[x]$ (see § 8.7), *the ring of power series* $P\{x\}$ (see § 8.8), and *the field of Laurent series* (see § 13.7). If $a \in R$ $(a \neq 0)$, then let $w(a)$ be the exponent of the least power of the indeterminate x that occurs in the expression for the polynomial (series) a with a non-zero coefficient; furthermore, let us set $w(0) = \infty$. *In this way,*

an integer-valued logarithmic valuation is introduced in R: the verification of the conditions 1_0), 2_0), $3_0'$) does not present any difficulty.

7. Returning to the general concept of a valuation (see § 44.1) we mention that condition 2) is sometimes replaced by the weaker condition

2′) $w(ab) \leqq w(a)w(b)$.

In this case we speak of a *pseudo-valuation.*

Important examples of rings with pseudo-valuations are certain rings of functions. For example, consider the ring of all continuous real functions $f(x)$ defined on the interval $[0, 1]$ of the real line. Setting

$$w(f) = \max_{x \in [0,\ 1]} |f(x)|,$$

we arrive at a real pseudo-valuation of this ring, because the conditions 1), 2′)—but not 2)—, and 3) are satisfied.

§ 45. Logarithmic Valuations of Fields

1. Logarithmic valuations are used mainly in the theory of fields and the theory of integral domains, their applications being based specifically on the concepts and results to be expounded in the present section.

Let w be a logarithmic valuation of a commutative field P (see § 44.6) with G as groupoid of values. From the results of § 44.4 and the definition of a logarithmic valuation it follows that G is an additively written, linearly ordered abelian group supplemented by the symbol ∞.

We denote by R_w the set of all those elements $a \in P$ for which $w(a) \geqq 0$. In virtue of 1_0), 2_0), and $3_0'$), R_w is a subring of P; it is called the *valuation ring* of w in P.

Since w determines a homomorphic mapping of the multiplicative group of P onto G, the unit element of P belongs to R_w.

Furthermore, if $a \in P$, then $w(a) = 0$ if and only if both a and a^{-1} belong to R_w, that is, if a is an invertible element of R_w, because

$$w(1) = 0. \tag{1}$$

2. *Let w and w' be two logarithmic valuations of a commutative field P with value groups G and G', respectively, both supplemented by the symbol ∞. Their valuation rings, regarded as subrings of P, coincide:*

$$R_w = R_{w'}. \tag{2}$$

if and only if there exists an isomorphic mapping φ of the ordered group G onto the ordered group G' such that for all non-zero elements $a \in P$

$$(w(a))\varphi = w'(a). \tag{3}$$

For suppose that (2) holds and that $w(a) = 0$. As mentioned above, a is then an invertible element of R_w and consequently of $R_{w'}$, so that $w'(a) = 0$. The converse statement is obvious.

Thus, the valuations w and w' determine homomorphisms of the multiplicative group of P having one and the same kernel. Therefore, by § 18.6, there exists an isomorphism φ of G onto G' satisfying the condition (3).

Finally, if $a \in G$ and $a \geqq 0$, then R_w contains an element a for which $w(a) = a$. It then follows from (2) that $w'(a) = a\varphi \geqq 0$ in G'. Thus, the isomorphism maps positive elements of G onto positive elements of G' and it therefore satisfies all the requirements of the theorem. The converse statement of the theorem is obvious. //

3. We have seen that a survey of all possible logarithmic valuations of a commutative field P reduces to a survey of all those subrings of the field that can serve as valuation rings.

A subring R of a commutative field P can be one of its valuation rings if and only if for every $a \in P$ $(a \neq 0)$ at least one of the elements a and a^{-1} belongs to R.

In one direction, the statement of the theorem is almost obvious: if $a \notin R_w$, i.e., $w(a) < 0$, then it follows from (1) that $w(a^{-1}) > 0$, and therefore $a^{-1} \in R_w$.

Now let R be a subring satisfying the conditions of the theorem. It therefore contains the unit element of P. The set S of invertible elements of R is obviously a subgroup of the multiplicative group P^* of P. We let G denote the factor group P^*/S, written additively.

Since R contains no divisors of zero, the non-zero elements of R form a sub-semigroup R^* of P^*. Since $S \subset R^*$, we can select the sub-semigroup $H = R^*/S$ in the group $G = P^*/S$. Let us show that H is the semigroup of positive elements for a suitable linear ordering of G, that is, let us verify the conditions 1)-4) of § 41.3.

Condition 1) (§ 41.3) follows from the fact that the zero element of G is the subgroup S contained in R^*. Next, if $a \in H$ and $-a \in H$, then R^* also contains the elements of the coset of a and their inverses, and then all these elements occur in S, i.e., a is the zero element of G; this proves condition 2). The validity of condition 3) is obvious. Finally, condition 4) follows from the assumption made about R in the statement of the theorem.

We now supplement the linearly ordered group G by the symbol ∞ with the properties (2) of § 44.6. When we set $w(0) = \infty$ and for every $a \in P$ $(a \neq 0)$ take as $w(a)$ that element of G which, as a coset of S, contains a, we obtain a logarithmic valuation of P.

For it is obvious that condition 1_0) holds. The validity of 2_0) in the case where both elements a and b are different from zero follows from the fact that the mapping $a \to w(a)$ $(a \neq 0)$ is a homomorphism of P^* onto G; if one of the elements a or b is equal to zero, then 2_0) follows from (2) of § 44.6.

Finally, let us verify condition $3_0'$). Since, by (1) and 2_0),

$$0 = w(1) = w(-1) + w(-1),$$

and since, by § 41.4, G does not contain non-zero elements of finite order, we have $w(-1) = 0$; and therefore, again by 2_0), for all $b \in P$,

$$w(-b) = w(b).$$

Therefore condition $3_0'$) is equivalent in our case to

$$w(a+b) \geqq \min(w(a), w(b)),$$

which we now proceed to prove.

The condition is obviously satisfied if at least one of the elements a, b is equal to zero. If they are both different from zero, then suppose that, say, $w(b) \leqq w(a)$, so that $a/b \in R$. Since the unit element of P is contained in R, we also have $1 + a/b \in R$, so that, in virtue of 2_0) and (1),

$$0 \leqq w(1 + a/b) = w[(a+b)b^{-1}] = \\ = w(a+b) + w(b^{-1}) = w(a+b) - w(b).$$

Hence

$$w(a+b) \geqq w(b) = \min(w(a), w(b)),$$

which completes the proof. //

4. If w is a logarithmic valuation of a field P with valuation ring R_w, then the elements $a \in P$ for which $w(a) > 0$ form an ideal I_w in R_w (see § 13.8), the so-called *valuation ideal* of w. For if $a, b \in I_w$, then it follows from $3_0'$) that $a - b \in I_w$; if $a \in I_w$, $x \in R_w$, i.e., $w(a) > 0$, $w(x) \geqq 0$, then by 2_0),

$$w(ax) = w(a) + w(x) > 0,$$

so that $ax \in I_w$.

Since, by § 45.1, every element $a \in R_w$, $a \notin I_w$ is invertible in R_w, the factor ring R_w/I_w is a field; it is called the *residue field* of w.

5. As an important example, let us find all the non-trivial logarithmic valuations of the field of rational numbers. Let w be such a valuation and let R_w and I_w be its valuation ring and valuation ideal, respectively. Since $1 \in R_w$, $1 \in I_w$, we see that R_w contains the whole ring of integers C, and the intersection $I_w \cap C$ is a proper ideal of C. This intersection is different from zero, because otherwise all non-zero integers would be invertible in R_w, i.e., this ring would coincide with the whole field of rational numbers, which is impossible for a non-trivial valuation. Therefore, by §§ 13.8 and 11.2, the ideal $I_w \cap C$ is the collection (p) of integers that are multiples of a certain natural number p $(p > 1)$.

Since the factor ring R_w/I_w is a field, the product of two integers not in (p) cannot belong to (p). This is, of course, equivalent to the statement that *p must be a prime number.*

Every integer n that is relatively prime to p does not belong to I_w and, by § 45.1, is therefore invertible in R_w, so that $n^{-1} \in R_w$. This shows that *R_w contains the whole ring R_p of those rational numbers whose denominators are prime to p.*

In fact, we have

$$R_w = R_p.$$

For if R_w contains a rational number m/n such that $(m, p) = 1$ and $n \in (p)$, then $m \notin I_w$, i.e., $m^{-1} \in R_w$, and then

$$n^{-1} = (m/n) \cdot m^{-1} = R_w,$$

that is, the number n proves to be invertible in R_w, in contradiction to the fact that $n \in I_w$.

Conversely, for every prime number p, the subring R_p of the field of rational numbers satisfies the conditions of the theorem of § 45.3. Thus, *the valuation rings of the field of rational numbers are exactly the rings R_p, where p ranges over all the prime numbers.* //

6. Let us now find the logarithmic valuation of the field of rational numbers for which R_p is the valuation ring.

Every rational number a $(a \neq 0)$ has a unique expression in the form

$$a = (m/n) \cdot p^k, \tag{4}$$

where m and n are prime to each other and to p, and k is an integer, positive, zero, or negative. We set

$$w(a) = k \qquad (a \neq 0) \tag{5}$$

and

$$w(0) = \infty.$$

We thus obtain a logarithmic valuation of the field of rational numbers with values in the ordered additive group of integers supplemented by the symbol ∞ : the verification of the conditions 1_0), 2_0), and $3_0'$) does not present any difficulties. This is called the *p-adic valuation* of the field of rational numbers.

From (4) and (5) it follows that *the valuation ring and the valuation ideal of the p-adic valuation are the ring R_p and its ideal R_p' consisting of those rational numbers whose numerators in the reduced form are divisible by p.* We note that

$$R_p = \{R_p', C\}, \tag{6}$$

where C is the ring of integers. For if $m/n \in R_p$, then it follows from $(n, p) = 1$ that the numbers sn ($s = 0, 1, 2, \ldots, p-1$) lie in distinct cosets of the ideal (p) in C. This is then also true of the numbers $m + sn$ ($s = 0, 1, 2, \ldots, p-1$); and therefore one of these numbers, for example $m + tn$, is divisible by p. Then

$$(m + tn)/n \in R_p',$$

but

$$m/n = [(m + tn)/n] - t.$$

Using (6), the fact that $R_p' \cap C = (p)$, and that R_p' is an ideal of R_p, we obtain from the isomorphism theorem (see § 20.2) that

$$R_p/R_p' \cong C/(p).$$

The residue field of the p-adic valuation is therefore isomorphic to the prime field C_p of characteristic p (see §§ 18.8 and 18.11).

7. Of course, the p-adic valuation of the field of rational numbers can also be regarded not as a logarithmic valuation, but as a non-archimedean valuation in the sense of § 44.3. In order to accomplish the transition inverse to the one described in § 44.6, it is sufficient to take an arbitrary real number r ($r > 1$) and, the rational number a being expressed in the form (4), to set

$$w(a) = r^{-k},$$

where furthermore $w(0) = 0$.

❖ For every archimedean valuation w of the field of rational numbers we can find a real number α ($0 < \alpha \leqq 1$) such that for all rational numbers a,

$$w(a) = |a|^{\alpha},$$

where $|a|$ is the absolute value of a [A. Ostrowski, *Acta Math.*, vol. 41 (1918), pp. 271-84]. ❖

8. Let w be a real valuation of commutative field P in the sense of § 44.1. A sequence of elements

$$a_1, a_2, \ldots, a_n, \ldots \tag{7}$$

of P is called a *fundamental sequence* if for every real number $\varepsilon > 0$ there exists a natural number $N = N(\varepsilon)$ such that

$$w(a_n - a_m) < \varepsilon \qquad \text{for} \qquad n, m > N.$$

The fundamental sequence (7) is called *convergent* if P contains an element a, the *limit* of the sequence, such that for every $\varepsilon > 0$ there exists a natural number $N = N(\varepsilon)$ satisfying the condition

$$w(a_n - a) < \varepsilon \qquad \text{for} \qquad n > N.$$

A convergent sequence having the limit 0 is called a *null sequence.*

A field P is called *complete* with respect to w if all of its fundamental sequences are convergent.

❖ For every commutative field P with a real valuation w there exists a *completion,* that is, a commutative field $\overline{P}$ with a real valuation $\overline{w}$, complete with respect to this valuation and such that, first, $P \subseteq \overline{P}$, second, $\overline{w}$ is an extension of w, and third, every element of $\overline{P}$ is the limit of a fundamental sequence of P.

If the valuation introduced in the field of rational numbers is the absolute value, then the completion of this field is the field of real numbers, also with the absolute value as valuation. But if the field of rational numbers is given a p-adic valuation, then its completion is the field of p-adic numbers (see § 19.12). Here the logarithmic value of a non-zero p-adic number written in the form (17) of § 19.12 is taken to be the integer k. ❖

9. Two fields P and P' with real valuations are called *isomorphic* if there exists an isomorphism between them for which null sequences of one field go over into null sequences of the other and vice versa.

❖ Every field with an archimedean real valuation is isomorphic to a subfield of the field of complex numbers with the modulus of complex numbers as valuation [A. Ostrowski, *Acta Math.*, vol. 41 (1918), pp. 271-84]. ❖

§ 46. Albert's Theorem on Normed Algebras

1. Suppose that in a commutative field P a real valuation is given, which we now denote by $|\alpha|$ $(\alpha \in P)$. Let R be an algebra over P and suppose that in R, as a ring, a real valuation w with the properties 1)-3) of § 44.1 is given. We call R a *normed algebra* if for arbitrary $a \in R$, $\alpha \in P$

$$w(\alpha a) = |\alpha| w(a). \tag{1}$$

This equation is also assumed to hold when we speak of a *pseudo-normed algebra* (see § 44.7).

2. *Every real finite-dimensional algebra R can be pseudo-normed* [A. A. Albert, *Ann. of Math.*, vol. 48 (1947), pp. 495-501].

For let a basis x_i $(i = 1, 2, \ldots, n)$ be chosen in R with the multiplication table

$$x_i x_j = \sum_{k=1}^{n} \varepsilon_{ij}^k x_k.$$

If μ is a non-zero real number, then the elements $y_i = \mu x_i$ $(i = 1, 2, \ldots, n)$ also form a basis with

$$y_i y_j = \sum_{k=1}^{n} \delta_{ij}^k y_k,$$

where

$$\delta_{ij}^k = \mu \varepsilon_{ij}^k.$$

As we can easily do, we choose the number μ so that we have, for all $i, j, k = 1, 2, \ldots, n$,

$$|\delta_{ij}^k| \leqq n^{-1}, \tag{2}$$

where $|\delta|$ is the absolute value of δ. When we now define $w(a)$, for every element $a \in R$ expressible in the form $a = \sum_{i=1}^{n} \alpha_i y_i$, by the equation

$$w(a) = \sum_{i=1}^{n} |\alpha_i|,$$

we introduce a pseudo-norm in R.

In fact, if

$$a = \sum_{i=1}^{n} \alpha_i y_i, \qquad b = \sum_{j=1}^{n} \beta_j y_j,$$

then

$$ab = \sum_{k=1}^{n} \left(\sum_{i,j=1}^{n} \delta_{ij}^k \alpha_i \beta_j \right) y_k,$$

and therefore, by (2),

$$w(ab) = \sum_{k=1}^{n} \left| \sum_{i,j=1}^{n} \delta_{ij}^k \alpha_i \beta_j \right| \leqq \sum_{i,j,k=1}^{n} |\delta_{ij}^k| |\alpha_i| |\beta_j| \leqq$$

$$\leqq \sum_{i,j=1}^{n} |\alpha_i| |\beta_j| = w(a) w(b).$$

This shows that condition 2') of § 44.7 holds, and obviously all the other conditions in the definition of a pseudo-normed algebra are satisfied. //

3. By way of contrast, there is the following *Theorem of Albert* [*Ann. of Math.*, vol. 48 (1947), pp. 495-501].

The field of real numbers, the field of complex numbers, the field of quaternions, and the Cayley algebra are the only finite-dimensional real normed algebras with a unit element.

We note that these four algebras actually satisfy all the conditions of the theorem (see § 44.1). On the other hand, *if a ring R has a real valuation w, then R has no divisors of zero*: if $a, b \in R$, $a \neq 0$, $b \neq 0$, $ab = 0$, then $w(a) \neq 0$, $w(b) \neq 0$, but $w(a)w(b) = w(ab) = 0$, which is impossible. Therefore, by the Generalized Frobenius Theorem (see § 38.1) we only have to prove the following statement:

Every finite-dimensional real normed algebra K having a unit element is alternative.

4. We begin the proof with a few remarks. First of all, we shall denote the norm in an algebra K by $|a|$, where $a \in K$.

Next, as we have mentioned above, K has no divisors of zero, so that by § 38.2, K is an algebra with unique division. From this it follows that for $a \neq 0$ ($a \in K$) the right multiplication by a (see § 31.6), being a linear transformation of the additive vector space of K, is non-singular. We denote this right multiplication by R_a; the identity linear transformation will be denoted by E.

Finally, since K has a unit element 1, by § 36.2 its center contains a subfield D isomorphic to the field of real numbers; this is the set of elements of the form $\alpha \cdot 1 = \alpha$, where α is a real number.

LEMMA 1: *If $a \notin D$, then the linear transformation R_a has no real eigenvalues.*

For if α is such an eigenvalue, then the linear transformation $R_a - \alpha E = R_{a-\alpha}$ is singular, so that, by what we have said above, $a - \alpha = 0$, and hence $a = \alpha \in D$.

5. LEMMA 2: *If $a \notin D$ and $|a| = \alpha$, then the modulus of every eigenvalue λ_0 of the linear transformation R_a is equal to α.*

For let

$$\lambda_0 = \varrho(\cos\varphi + i \sin\varphi). \tag{3}$$

The linear transformation $R_a - \lambda_0 E$ of the complex vector space that is an extension of the real additive vector space of K must be singular. Therefore K contains elements b and c such that

$$b + ci \neq 0,$$

and hence

$$\delta = |b| + |c| > 0. \tag{4}$$

and

$$(b + ci)(R_a - \lambda_0 E) = 0.$$

Therefore

$$(b + ci)R_a = \lambda_0(b + ci),$$

and consequently

$$(b + ci)R_a^m = \lambda_0^m(b + ci) \tag{5}$$

for all natural numbers m. Since R_a is a linear transformation of the real vector space K, (5) and (3) yield the equations

$$\begin{aligned} bR_a^m &= \varrho^m(b \cos m\varphi - c \sin m\varphi), \\ cR_a^m &= \varrho^m(b \sin m\varphi + c \cos m\varphi). \end{aligned} \tag{6}$$

Using the definition of a normed algebra—in particular, (1) and (4) and the inequalities $|\cos m\varphi| \leqq 1$, $|\sin m\varphi| \leqq 1$—we obtain from (6)

$$\begin{aligned} |b|\alpha^m &= |bR_a^m| \leqq \varrho^m\delta, \\ |c|\alpha^m &= |cR_a^m| \leqq \varrho^m\delta. \end{aligned}$$

Hence, by (4), $\delta\alpha^m \leqq 2\varrho^m\delta$, or $(\alpha\varrho^{-1})^m \leqq 2$; and since m is arbitrary, $\alpha \leqq \varrho$.

On the other hand, from (6) we derive the equations

$$\begin{aligned} (b \cos m\varphi + c \sin m\varphi)R_a^m &= \varrho^m b, \\ (c \cos m\varphi - b \sin m\varphi)R_a^m &= \varrho^m c, \end{aligned}$$

which lead to the inequalities

$$\varrho^m|b| \leqq \delta\alpha^m, \qquad \varrho^m|c| \leqq \delta\alpha^m.$$

Hence $\varrho^m\delta \leqq 2\delta\alpha^m$, or $(\varrho\alpha^{-1})^m \leqq 2$; and therefore $\varrho \leqq \alpha$. Hence we see that $\varrho = \alpha$, as required.

6. LEMMA 3: *If $a \notin D$, then the characteristic polynomial of the linear transformation R_a is a power of an irreducible quadratic trinomial.*

For if λ_0 is an eigenvalue of R_a and $b = a + 1$, then $R_b = R_a + E$, and therefore the number $\lambda_0 + 1$ is an eigenvalue for R_b. Let $\alpha + \beta i$, $\gamma + \delta i$ be two eigenvalues of R_a. It follows from Lemma 2 that these eigenvalues have equal moduli, i.e.,

$$\alpha^2 + \beta^2 = \gamma^2 + \delta^2.$$

Applying the same argument to the eigenvalues of R_b—clearly, $b \notin D$—we obtain

$$(\alpha + 1)^2 + \beta^2 = (\gamma + 1)^2 + \delta^2.$$

Hence $2\alpha = 2\gamma$, or $\alpha = \gamma$, and therefore $\beta^2 = \delta^2$, that is, $\beta = \pm \delta$. These results, in conjunction with Lemma 1, conclude the proof of the lemma.

7. LEMMA 4: *If $a \notin D$, then the minimal polynomial of the linear transformation R_a is an irreducible quadratic trinomial.*

By Lemma 3, the characteristic polynomial of R_a has the form

$$\chi(\lambda) = (\lambda^2 - a\lambda - \beta)^h,$$

where

$$\gamma = -a^2/4 - \beta > 0.$$

Therefore, the minimal polynomial is

$$q(\lambda) = (\lambda^2 - a\lambda - \beta)^k \qquad (k \leqq h),$$

so that

$$(R_a^2 - aR_a - \beta E)^k = 0. \tag{7}$$

Let $k > 1$. If

$$b = \gamma^{-1/2}(a - a/2),$$

so that $b \notin D$, then the equations

$$(R_a^2 - aR_a - \beta E)^l = 0 \quad \text{and} \quad (R_b^2 + E)^l = 0$$

are equivalent. Thus,

$$(R_b^2 + E)^k = 0, \qquad (R_b^2 + E)^{k-1} \neq 0. \tag{8}$$

Since i is an eigenvalue of R_b, we have $|b| = 1$, by Lemma 2. Setting $S = R_b^2$, we now obtain for every $c \in K$

$$|cS| = |(cb)b| = |c|.$$

Hence

$$|cS^m| = |c| \tag{9}$$

for all $c \in K$ and all natural numbers m.

If $T = (S + E)^{k-2}$—this is where we make use of the fact that $k \geqq 2$—then by (8),

$$T \neq 0, \quad T(S + E) \neq 0, \quad T(S + E)^2 = 0.$$

Therefore there exists an element c' such that

$$c'(S+E)^{k-1} \neq 0,$$

i.e.,

$$c = c'T \neq 0, \quad c(S+E) \neq 0, \quad c(S+E)^2 = 0. \tag{10}$$

From the last equation it follows that

$$cS^2 = -c(2S+E),$$

i.e., for $m = 2$ we have the equation

$$cS^m = (-1)^{m-1}c[mS + (m-1)E]. \tag{11}$$

If the equation (11) is already proved for a given m, then

$$cS^{m+1} = (-1)^{m-1}c[mS^2 + (m-1)S] =$$
$$= (-1)^m c[2mS + mE - (m-1)S] = (-1)^m c[(m+1)S + mE],$$

i.e., (11) is true for all m. By a slight change in notation and by taking norms on both sides, we see, in virtue of (9), that

$$|c| = |mc(S+E) - c|. \tag{12}$$

If

$$|c(S+E)| = \delta,$$

then, by (10), $\delta > 0$. However,

$$|mc(S+E) - c| \geqq m\delta - |c|,$$

so that by (12), we have the inequality

$$2|c| \geqq m\delta$$

for all m, which is impossible. This shows that $k = 1$, and Lemma 4 is proved.

8. By virtue of (7), this lemma yields the equation

$$R_a^2 = aR_a + \beta E,$$

so that, since K has a unit element,

$$a^2 = aa + \beta.$$

Hence, for every $b \in K$,

$$(ba)a = bR_a^2 = b(aR_a + \beta E) = aba + \beta b = b(aa + \beta) = ba^2.$$

Thus, the equation $(ba)a = b(aa)$ is proved for $a \notin D$; however, it holds equally well for $a \in D$, because D is contained in the center of K. The equation $a(ab) = (aa)b$ is verified similarly. Hence, by the theorem of § 37.3, K is an alternative algebra, and so we have proved Albert's Theorem. //

❖ The field of real numbers, the field of complex numbers, the field of quaternions, and the Cayley algebra are the only real normed division algebras with a unit element; it need not be assumed beforehand that these algebras are finite-dimensional [A. A. Albert, *Bull. Amer. Math. Soc.*, vol. 55 (1949), pp. 763-68; F. B. Wright, *Proc. Nat. Acad. USA*, vol. 39 (1953) pp. 330-32].

The field of real numbers, the field of complex numbers, and the field of quaternions are the only pseudo-normed real associative algebras that do not contain a non-zero element x with the following property: There exists a sequence y_n, not itself a null sequence (see § 45.8), such that at least one of the sequences xy_n and $y_n x$ is a null sequence [I. Kaplansky, *Duke Math. J.*, vol. 16 (1949), pp. 399-418].

The field of complex numbers is the only pseudo-normed complex associative division algebra [S. Mazur, *C. R. Paris*, vol. 207 (1938), pp. 1025-27; I. M. Gel'fand, *Mat. Sb.*, vol. 9 (1941), pp. 3-24]. ❖

§ 47. Closure. Topological Spaces

1. We say that a *closure relation* is given in a partially ordered set M if there is assigned to every element $a \in M$ a uniquely determined element $\bar{a} \in M$, the so-called *closure* of a, where for all $a, b \in M$ the following conditions are satisfied:

1_0) $a \leqq \bar{a}$;

2_0) If $a \leqq b$, then $\bar{a} \leqq \bar{b}$;

3_0) $\bar{\bar{a}} = \bar{a}$; that is, the closure of an arbitrary element is its own closure.

An element a is called *closed* if it coincides with its closure. By 3_0), the closures of all elements are closed and, obviously, they constitute the only closed elements. Furthermore, it follows from 2_0) that $\bar{a}$ is contained in every closed element containing a, so that *a closure relation is uniquely determined when the system of closed elements is assigned.*

In every partially ordered set we have the *trivial closure* if we put $\bar{a} = a$ for all a.

2. A closure relation can be introduced, in particular, in the system $\tilde{M}$ of all subsets of an arbitrary set M, partially ordered by set-theoretical inclusion. Henceforth, it will be in just this sense that we shall speak of the

closure relation given in an arbitrary set M. Clearly, the meaning of such concepts as *the closure of a subset* of M and *closed subset* is well defined; also, the statements of the conditions 1_0)-3_0) of § 47.1 remain valid in essence.

Observe that if a closure relation is given in a set M, then *the set M itself is closed,* since it is the closure of itself. Furthermore, *the intersection of an arbitrary system of closed subsets of M is itself closed.*

For let A_α be a system of closed subsets of M (α ranges over an index set) and

$$B = \bigcap_\alpha A_\alpha.$$

Then, by condition 2_0) of § 47.1, it follows from $B \subseteqq A_\alpha$ that $\overline{B} \subseteqq \overline{A}_\alpha = A_\alpha$, so that

$$\overline{B} \subseteqq \bigcap_\alpha A_\alpha = B.$$

On the other hand, $B \subseteqq \overline{B}$ by condition 1_0) of § 47.1, and therefore we have $\overline{B} = B$. //

The converse theorem also holds:

In every set M a closure relation can be assigned by taking as the system of closed subsets an arbitrary system Σ of subsets that contains the intersection of each of the subsystems of Σ and that also contains M itself.

For if we take as the closure $\overline{A}$ of an arbitrary subset A of M the intersection of all subsets of Σ containing A (such subsets exist: for example, M itself), then the conditions 1_0)-3_0) are satisfied and the closed subsets are precisely those of Σ. //

3. Two partially ordered sets with closures are naturally called *isomorphic* if there exists an isomorphic correspondence in the sense of § 4.3 between them under which images and inverse images of closed elements are closed. In this sense, we can speak of an *isomorphic embedding* of a partially ordered set with closure in another.

❖ If M is a partially ordered set with closure, then a closure relation can be assigned in the partially ordered system $\tilde{M}$ of all subsets of M such that M becomes isomorphically embedded in $\tilde{M}$. ❖

4. The concept of a set with a closure relation for subsets, as introduced in § 47.2, is far too wide. By imposing certain additional restrictions on the closure relation, we arrive at the concept of a topological space, one of the most important of mathematical concepts, which in many cases plays a role paralleling that of partially ordered sets.

We say that a *topology* is given in a set M, or that M is a *topological space,* if a closure is assigned in M that satisfies, in addition to conditions

1_0-3_0) (or the equivalent conditions stated in the theorem of § 47.2), the following additional conditions as well:

4_0) The closure of the union of two (and therefore of an arbitrary finite number) of subsets of M is equal to the union of the closures of these subsets:

$$\overline{A \cup B} = \overline{A} \cup \overline{B};$$

5_0) Every subset consisting of a single element is closed.

Actually, *the conditions* 1_0) *and* 2_0) *can be deduced from* 4_0) *and* 5_0). For let A be an arbitrary subset of M and $a \in A$. Then by the conditions 4_0) and 5_0),

$$\overline{A} = \overline{A \cup a} = \overline{A} \cup \overline{a} = \overline{A} \cup a,$$

i.e., $a \in \overline{A}$, and this implies the inclusion $A \subseteqq \overline{A}$, thus establishing the condition 1_0). Furthermore, if the subsets A and B are such that $A \subseteqq B$, then by condition 4_0),

$$\overline{B} = \overline{A \cup B} = \overline{A} \cup \overline{B},$$

so that $\overline{A} \subseteqq \overline{B}$; and this proves the condition 2_0).

If a topological space M contains more than one element—and only such spaces are of interest—*then its empty subset is closed.* For it is closed as the intersection of any two distinct elements of M, which, by condition 5_0), are closed subsets.

Finally, we note the following consequence of condition 4_0): *The union of an arbitrary finite number of closed subsets of a topological space is closed.*

5. The system of all closed subsets of a topological space M is partially ordered by inclusion. Partially ordered by inclusion, also, is the system of all complements in M of closed subsets—that is, the system of *open subsets.* We obtain an inversely isomorphic correspondence (see § 4.6) between these two partially ordered sets if we assign to every closed subset A the open subset $M \setminus A$. This enables us to derive new results concerning open sets by dualization from results already known concerning closed sets.

For example, bearing in mind that the complement of the intersection of a given system of subsets of an arbitrary set is always the union of the complements of these subsets (and the same with 'intersection' and 'union' interchanged), we see that *the union of an arbitrary system of open subsets and the intersection of an arbitrary finite number of open subsets is itself open. The space M and the empty subset are open,* because each is the complement of the other.

6. Every closure relation in a set M, and therefore every topology, is, as we know, completely determined when the systems of all closed subsets are given. A topology in M can therefore also be defined by giving the system of all open subsets. Keeping in mind that the union of an arbitrary system of open sets is open, we can confine ourselves to giving a system of open subsets Σ such that every (non-empty) open subset of M is a union of subsets in Σ. Such a system Σ, a so-called *complete system of neighborhoods* of M, is of course not uniquely determined.

The subsets that form a given complete system of neighborhoods Σ are called *neighborhoods*; all the neighborhoods of Σ that contain a given element $a \in M$ form a *complete system of neighborhoods of this element.*

Let a topology in M be given by means of a complete system of neighborhoods Σ. Then *a subset A of M is open if and only if for every $a \in A$ we can find a neighborhood U in Σ such that $a \in U$ and $U \subseteq A$;* this follows from the fact that A is open if and only if it is a union of neighborhoods from Σ.

If A is an arbitrary subset of M, then its closure $\overline{A}$ consists of precisely those elements every neighborhood of which contains at least one element of A.

For if an element $x \notin \overline{A}$, then x is contained in the open set $M \setminus \overline{A}$, and hence there exists a neighborhood of x contained in $M \setminus \overline{A}$, i.e., having an empty intersection with A. On the other hand, if an element x has a neighborhood U not containing elements of A, then A is contained in the closed set $M \setminus U$, so that by property 2_0) of § 47.1,

$$\overline{A} \subseteq \overline{M \setminus U} = M \setminus U,$$

i.e., $x \notin \overline{A}$. //

From this it follows that *a subset A of a topological space M is closed if and only if every element x each neighborhood of which contains at least one element of A itself belongs to A.*

7. *A system Σ of subsets of a set M is a complete system of neighborhoods for a topology given in M if and only if*

α) for every ordered pair of distinct elements a, b of M we can find a subset U belonging to Σ such that $a \in U$, $b \notin U$;

β) for any two subsets U and V occurring in Σ and containing an element a we can always find in a subset W in Σ such that $a \in W$ and $W \subseteq U \cap V$.

For let Σ be a complete system of neighborhoods of the topological space M. If a and b are distinct elements of M, then, b being closed, the subset $M \setminus b$ is open and is therefore the union of neighborhoods from Σ. Since $a \in M \setminus b$, we can find a neighborhood U in Σ, by § 47.6, such that $a \in U$ and

$U \subseteq M \setminus b$, so that $b \notin U$. This proves α). On the other hand, if U and V are neighborhoods of a, then by § 47.5, their intersection $U \cap V$ is open, and since a is contained in this intersection, we see, again by § 47.6, that there exists a neighborhood W of a contained in $U \cap V$.

Now let Σ be an arbitrary system of subsets of M having the properties α) and β). For every subset $A \subseteq M$ we denote by $\overline{A}$ the collection of those elements $x \in M$ for which every subset from Σ containing x contains at least one element of A. We shall show that a topology is introduced in M if we take this $\overline{A}$ as closure. By § 47.4, it is sufficient to show that the conditions 3_0), 4_0), and 5_0) are satisfied.

If $x \in \overline{\overline{A}}$, then every subset U in Σ containing x contains at least one element $y \in \overline{A}$, so that, again by the definition of $\overline{A}$, U contains at least one element of A. Hence it follows that $x \in \overline{A}$, i.e., $\overline{\overline{A}} \subseteq \overline{A}$. On the other hand, since it is obvious that $A \subseteq \overline{A}$, we see that $\overline{A} \subseteq \overline{\overline{A}}$ and thus $\overline{\overline{A}} = \overline{A}$; that is, condition 3_0) is proved.

Next, if A and B are arbitrary subsets of M and x is an arbitrary element of the union $\overline{A} \cup \overline{B}$, then every subset from Σ containing x contains at least one element either of A, or of B, i.e., at least one element of $A \cup B$, so that $x \in \overline{A \cup B}$. This shows that $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$. On the other hand, if $y \in \overline{A \cup B}$, then every subset from Σ containing y contains at least one element of $A \cup B$. Suppose that there exist in Σ two subsets U and V each containing y such that U does not contain elements of A, nor V elements of B. Then, by condition β), we can find in Σ a subset W containing y and lying in the intersection $U \cap V$, i.e., not containing elements of A or of B, which is impossible. Thus, either all subsets in Σ containing y contain elements of A, and then $y \in \overline{A}$, or they all contain elements of B, and then $y \in \overline{B}$. Hence it follows that $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$. This shows that condition 4_0) is valid.

Finally, the validity of 5_0) follows from the fact that for every element $b \in M$ its closure $\overline{b}$ consists of b only, because by condition α) for every other element a we can find a subset in Σ that contains a but does not contain b.

To complete the proof of the theorem, it remains to show that the system Σ is a complete system of neighborhoods for the topology that we have constructed in M. If U is a subset in Σ, then its complement $M \setminus U$ is closed, because every element of M not in $M \setminus U$ is contained in U, i.e., lies in a subset from Σ containing no element of $M \setminus U$. This proves that U is open. On the other hand, let A be an arbitrary subset and a an element of A. Then $M \setminus A$ is closed, $a \notin M \setminus A$, and therefore there exists a subset U in Σ containing a whose intersection with $M \setminus A$ is empty, so that $U \subseteq A$. The subset A can therefore be represented as a union of a system of subsets in Σ. //

8. Two sets with closures, M and M', are called *isomorphic* if there exists a one-to-one correspondence χ between them preserving the closure relation; that is, such that for all $A \subseteq M$,

$$\overline{A}\chi = \overline{A\chi}.$$

The mapping χ is then called an *isomorphism*. The inverse mapping χ^{-1} is, of course, also an isomorphism.

The relation of this concept to the concept of an isomorphism of partially ordered sets with closures introduced in § 47.3 is obvious. In the case of topological spaces, the term 'isomorphism' is usually replaced by 'homeomorphism'; we shall not, however, make use of this special term.

Under an isomorphism between sets with closures, closed subsets go over into closed subsets and vice versa, and by § 47.1, *this property can be taken as the definition of isomorphism.* Hence it follows that *in the case of topological spaces, an isomorphism can also be defined as a one-to-one correspondence between the spaces under which open subsets go over into open subsets and vice versa.*

If topologies are given in two topological spaces M and M' by means of complete systems of neighborhoods Σ and Σ', respectively, then a one-to-one mapping χ of M onto M' is an isomorphism if and only if for every element $a' \in M'$ and every neighborhood U' of M' there exists a neighborhood U of the element $a'\chi^{-1} \in M$ such that $U\chi \subseteq U'$ and for every element $a \in M$ and every neighborhood V of M there exists a neighborhood V' of $a\chi$ such that $V'\chi^{-1} \subseteq V$.

For suppose that χ is an isomorphism and that, for example, an element $a' \in M'$ and its neighborhood U' are given. Then the set $U'\chi^{-1} \subseteq M$ is open, so that it contains a neighborhood U of $a'\chi^{-1}$. It is clear that $U\chi \subseteq U'$.

Conversely, suppose that a mapping χ satisfies the conditions laid down in the theorem and that, for example, A is an open subset of M. If $a\chi$ is an arbitrary element of $A\chi$, then $a \in A$ and A contains a neighborhood V of a. By the condition of the theorem, a neighborhood V' of $a\chi$ then exists such that $V'\chi^{-1} \subseteq V$. Hence

$$V' = V'\chi^{-1}\chi \subseteq V\chi \subseteq A\chi.$$

This shows that $A\chi$ is an open subset of M'. //

When we apply this theorem to the case where M and M' coincide and χ is the identity mapping of M onto itself, i.e., $a\chi = a$ for all $a \in M$, we arrive at the condition for the *equivalence* of two complete systems of neighbor-

hoods Σ and Σ' given in M, that is, the condition under which these two complete systems of neighborhoods define the *same* topology in M: *this is the case if and only if every neighborhood of an arbitrary element* $a \in M$ *in one of the systems* Σ, Σ' *contains a neighborhood of* a *in the other system.*

9. A topology given in a set M induces a topology in every subset $A \subseteq M$. Namely, if the topology in M is given by a complete system of neighborhoods Σ, then we denote by Σ_A the set of all intersections $U \cap A$, where $U \in \Sigma$. Since the conditions α) and β) of the theorem of § 47.7 are valid, it follows immediately that the corresponding conditions hold for the system Σ_A, so that this system is a complete system of neighborhoods for a topology given in A.

If Σ' is another complete system of neighborhoods in M, equivalent to Σ, then the equivalence of the systems of neighborhoods Σ_A and Σ'_A in A can be verified without difficulty on the basis of § 47.8. Thus, the topology in M uniquely determines a topology in A. The topological space A obtained in this way is called a *subspace* of M.

From the definition of a subspace it follows immediately that *the open and the closed subsets of the subspace* A *of* M *are precisely the intersections with* A *of the open and the closed subsets, respectively, of* M.

10. The trivial closure (see § 47.1) introduced in the system of all subsets of an arbitrary set M obviously satisfies the conditions 4_0) and 5_0) of § 47.4, so that it is a topology. This is called the *discrete* topology; all the subsets of M in this topology are both open and closed. Every set can therefore be regarded as a discrete topological space; and in finite sets, where every subset is a union of a finite number of elements, only the discrete topology is possible.

As our first example of a non-discrete topological space, we have the *straight line*. If we regard it as the real line and for every subset A define the closure $\overline{A}$ as the set of numbers that are limits of convergent sequences of numbers from A, we obtain a topology: the conditions 3_0), 4_0), and 5_0) are easily verified. One of the complete systems of neighborhoods for this natural topology of the line is the system of all (open) intervals. Henceforth, in speaking of the real line as a topological space, we shall always mean its natural topology.

The plane and, more generally, every n-dimensional real euclidean space are also often regarded as topological spaces. A topology can be introduced in them in the same way as was done above for the line, by using coordinate-wise convergence. In the case of a plane, a complete system of neighborhoods is the system of all (open) discs; an equivalent complete system of neighborhoods is the system of all (open) squares.

§ 48. Special Types of Topological Spaces

1. As we know (see, for example, § 47.7), for any two distinct elements a and b of a topological space we can find an open subset U containing a but not b. The imposition of stronger conditions than this enable us to single out certain special important types of topological spaces.

For example, a topological space M is called a *Hausdorff* space if for any two distinct elements a and b of M we can find open subsets U and V such that $a \in U$, $b \in V$, and the intersection $U \cap V$ is empty.

If a topology in M is given by the complete system of neighborhoods Σ, then M is a Hausdorff space if and only if the intersection of the closures of all neighborhoods of an arbitrary element a consists of this element only.

For if M is a Hausdorff space and b is different from a, then by taking a neighborhood U of a and a neighborhood V of b such that $U \cap V$ is empty, we see that $b \notin \overline{U}$. On the other hand, if M satisfies the condition of the theorem and if a and b are two distinct elements of M, then there exists a neighborhood U of a such that $b \notin \overline{U}$. Then b is contained in the open subset $V = M \setminus U$, and $U \cap V$ is empty. //

2. A topological space is called *regular* if for every element a and every closed subset B not containing a we can find disjoint open subsets U and V such that $a \in U$, $B \subseteq V$.

Clearly, every regular space is a Hausdorff space. All the spaces mentioned in § 47.10 are regular.

If a topology in M is given by a complete system of neighborhoods Σ, then M is a regular space if and only if for every neighborhood U of an arbitrary element a we can find a neighborhood V of this element whose closure $\overline{V}$ is contained in U.

For if M is a regular space and U is a neighborhood of an element a, then we can find for a and for the closed subset $M \setminus U$ disjoint open subsets V and W such that $a \in V$, $(M \setminus U) \subseteq W$. Then V contains a neighborhood V_0 of a with

$$\overline{V}_0 \subseteq \overline{V} \subseteq M \setminus W \subseteq M \setminus (M \setminus U) = U.$$

Conversely, suppose that M has the property of the theorem and that an element a and a closed subset B with $a \notin B$ are given in M. By hypothesis, there exists then a neighborhood U of a whose closure $\overline{U}$ is contained in the open set $M \setminus B$. The intersection of the open subsets U and $M \setminus \overline{U}$ is empty, and since

$$M \setminus \overline{U} \supseteq M \setminus (M \setminus B) = B,$$

we have proved that M is regular. //

It is easy to verify that *the property of being a Hausdorff space or of being a regular space carries over to subspaces of the space.*

More special than the concept of a regular space is that of a *normal* space: this is a topological space in which for any two disjoint closed sets A and B there exist disjoint open sets U and V such that $A \subseteq U$, $B \subseteq V$.

❖ An intermediate class between the regular and the normal spaces are the *completely regular* topological spaces. These are spaces with the following property: For every element a and every closed set B not containing a, we can define a continuous real function f on the space (that is, continuous in the sense of the topology of the space) such that $f(a) = 0$, $f(b) = 1$ for all $b \in B$ and $0 \leqq f(x) \leqq 1$ for all x. ❖

3. A very important class of topological spaces are the compact spaces. A space M is called *compact* if it satisfies any of the following conditions, whose equivalence follows immediately from what we have said in § 47.5:

I) Given a system of open subsets of M such that M is the union of the subsets of this system, we can choose a finite subsystem such that M is the union of the subsets of this finite subsystem.

II) From every system of closed subsets whose intersection is empty we can choose a finite subsystem such that the intersection of these subsets is empty.

The so-called Heine-Borel Theorem, which is proved in courses in mathematical analysis, shows that every finite closed interval of the real line is a compact space.

❖ Each of the following two conditions is equivalent to I) and II) and can therefore be used to define a compact space:

III) The union of every chain (see § 4.1) formed from open subsets other than M is itself different from M.

IV) The intersection of every chain formed from non-empty closed subsets is itself non-empty. ❖

4. A topological space is called *locally compact* if every element of the space is contained in an open subset whose closure is compact.

Every compact space is locally compact. An example of a locally compact, but not compact, space is the real line. For, every point of the real line has as a neighborhood a finite open interval. The closure of such a neighborhood is a closed interval and so, as was mentioned above, is compact. On the other hand, the line itself is not compact: it is the union of all its intervals of finite length but cannot be represented as the union of a finite number of such intervals.

As another example of a locally compact, but not compact, space we can take any infinite discrete space.

5. *A closed subset of a compact space is compact and of a locally compact space is locally compact.*

We note, first of all, that if A is a closed subset of an arbitrary topological space M, then every subset B of A that is closed in A is also closed in M. For if $\overline{B}$ is the closure of B in M, then, as follows from § 47.9, $\overline{B} \cap A$ is the closure of B in A, and since B is closed in A,

$$\overline{B} \cap A = B. \tag{1}$$

However, as A itself is closed in M, it follows from $B \subseteq A$ that $\overline{B} \subseteq A$, so that, by (1), $\overline{B} = B$.

It is now clear that when condition II) of § 48.3 holds in a compact space M, then it also holds in every closed subspace of M. But if M is locally compact and A is a closed subspace of M, then for every $a \in A$ there exists a neighborhood U in M whose closure $\overline{U}$ is compact. Then the intersection $U \cap A$ is a neighborhood of a in A, and its closure in A is $\overline{U} \cap A$. This intersection, however, is compact, because it is closed in M (as the intersection of closed subsets) and is contained in the compact space $\overline{U}$. //

❖ Every open subset of a locally compact space—and, in particular, a compact space—is locally compact.

Every locally compact space is isomorphic to an open subspace of some compact space [P. S. Alexandrov, *Math. Ann.*, vol. 92 (1924), pp. 294-301]. ❖

6. A topological space M is called *disconnected* if it can be split into two non-empty disjoint closed subsets (each of which is obviously open as well); otherwise, it is called *connected*.

A topological space M is called *totally disconnected* if for any two distinct elements $a, b \in M$ there exists a decomposition of M into two disjoint closed subsets A and B such that $a \in A$, $b \in B$.

Every discrete space is, of course, totally disconnected. An example of a non-discrete totally disconnected space is the *rational line*, i.e., the subspace of the real line formed by the rational numbers. This is not discrete, because every neighborhood on the real line contains an infinite number of rational numbers. On the other hand, if a and b are two distinct rational numbers ($a < b$), then there exists an irrational number α between them, and we obtain the required decomposition by taking for A the set of rational numbers that are less than α and for B the set of rational numbers greater than α.

We note that every subspace of a totally disconnected topological space is itself totally disconnected.

§ 49. Topological Groups

1. An examination of various examples of algebraic structures that are at the same time topological spaces (see § 47.4) leads to the following definitions.

A groupoid G that is at the same time a topological space is called a *topological groupoid* if the multiplication in G is *continuous* in the given topology, that is, if for any two elements $a, b \in G$ and any neighborhood W of ab (see § 47.6) we can find neighborhoods U and V of a and b, respectively, such that $UV \subseteq W$, where UV is the set of elements of G that can be represented in at least one way as a product uv $(u \in U, v \in V)$.

A special case of a topological groupoid is a *topological semigroup*.

A group G is called a *topological group* if it is a topological semigroup and if, in addition, the operation of taking the inverse element is *continuous* in the given topology, that is, if for every element $a \in G$ and every neighborhood V of a^{-1} there exists a neighborhood U of a such that $U^{-1} \subseteq V$, where U^{-1} is the set of elements of the form u^{-1} for all $u \in U$.

It is easy to verify, by applying the condition of equivalence for complete systems of neighborhoods in a topological space (see § 47.8) that the continuity of the operation of multiplication and the continuity of the operation of going over to the inverse element are properties that do not depend on the choice of a complete system of neighborhoods in G. The intuitive meaning of continuity of the operations is that small changes in the factors (or in the element) correspond to small changes in the product (or in the inverse element).

If we introduce a topology in the group G that makes G into a topological group, we call this the *topologization* of G.

❖ There exist groups that are topological semigroups but not topological groups.

Every compact Hausdorff topological semigroup (see §§ 48.1 and 48.3) with the cancellation law (see § 11.1) is a group, and even a topological group [see R. Arens, *Bull. Amer. Math. Soc.*, vol. 53 (1947), pp. 623-30]. ❖

2. Every group can be regarded as a topological group with the discrete topology (see § 47.10). Of course, a finite group admits only the discrete topology. The problem of whether every infinite group admits a non-discrete topology is still open.

❖ Every infinite abelian group can be made into a topological group with a non-discrete topology [A. Kertesz and T. Szele, *Publ. Math.*, vol. 3 (1953), pp. 187-89]. ❖

3. *If A is an open subset of a topological group G and g is an element of G, then the sets Ag, gA, and A^{-1} are also open.*

We shall only prove the first of these statements. If $b = ag$ $(a \in A)$, then $a = bg^{-1}$. From the continuity of multiplication it follows that there exists a neighborhood U of b such that $Ug^{-1} \subseteq A$ (we recall that A is an open set). Hence $U \subseteq Ag$, that is, if the set Ag contains b it also contains a neighborhood of b. //

If g is an element of a topological group G, then the mappings (for all $x \in G$)

$$x \to xg, \quad x \to gx, \quad x \to x^{-1}$$

are isomorphisms (see § 47.8) *of G as a topological space onto itself.*

The fact that these mappings are one-to-one is obvious, and we can therefore refer to the preceding theorem, bearing in mind that the inverse mapping of the right *translation* $x \to xg$ is the right translation $x \to xg^{-1}$ and that the mapping $x \to x^{-1}$ is its own inverse. //

Hence it follows, by § 47.8, that *if A is a closed subset of a topological group G and g is an element of G, then the sets Ag, gA, and A^{-1} are also closed.*

4. If a and b are elements of a topological group G, then the translation $x \to xa^{-1}b$ $(x \in G)$ carries a into b. From this, and from what was established above, it follows that in proving properties of the space of a topological group concerning individual elements it is sufficient to examine only a fixed element—for example, the unit element. It also follows that for the purpose of giving a topology in a group it is not necessary to give an entire complete system of neighborhoods, but we can confine ourselves to a *complete system of neighborhoods of the unit element* (see § 47.6).

❖ A system Σ of subsets of a group G can be the complete system of neighborhoods of the unit element for a suitable topologization of the group if and only if the following conditions are satisfied:

1) The intersection of all the sets in Σ consists of the unit element of G only;

2) The intersection of any two sets of Σ contains a third set belonging to Σ;

3) For every set U of Σ we can find a set V in Σ such that $VV^{-1} \subseteq U$;

4) For every set U of Σ and every element $a \in U$ we can find a set V in Σ such that $Va \subseteq U$;

5) For every set U of Σ and every element a of G we can find a set V in Σ such that $a^{-1}Va \subseteq U$. ❖

5. *The space of every topological group G is a regular, and therefore a Hausdorff space* (see §§ 48.1 and 48.2).

Using the criterion for regularity given in § 48.2 and bearing in mind what was said in the preceding subsection, we choose an arbitrary neighborhood U of the unit element of G. Since $1 \cdot 1^{-1} = 1$, there exist neighborhoods V_1 and V_2 of the elements 1 and 1^{-1} ($= 1$, of course) such that $V_1 V_2 \subseteq U$. Further, there exists a neighborhood V_3 of the unit element such that $V_3^{-1} \subseteq V_2$. Finally, there exists a neighborhood V of the unit element contained in the intersection $V_1 \cap V_3$ (see § 47.7). It is clear that

$$VV^{-1} \subseteq U. \tag{1}$$

Let $a \in \overline{V}$. Since the set aV is open, by § 49.3, and since it contains the element $a = a \cdot 1$ of the closure of V, by § 47.6 we can find an element b of V in aV. Let $b = ac$ ($c \in V$). Hence, by (1),

$$a = bc^{-1} \in VV^{-1} \subseteq U,$$

i.e., $\overline{V} \subseteq U$. //

❖ The space of every topological group is completely regular (see § 48.2).

There are topological groups whose spaces are not normal [A. A. Markov, *Izv. AN SSSR, ser. matem.*, vol. 9 (1945), pp. 3-64]. ❖

6. Every subgroup A of a topological group G is, of course, itself a topological group in the topology that is induced in it (see § 47.9) by the topology of G.

The closure $\overline{A}$ of a subgroup A of a topological group G is itself a subgroup. If A is normal in G, then so is $\overline{A}$. If A is abelian, then so is $\overline{A}$.

For if $x, y \in \overline{A}$ and U is an arbitrary neighborhood of xy, then there exist neighborhoods V of x and W of y such that $VW \subseteq U$. Further, there exist elements $x', y' \in A$ such that $x' \in V$, $y' \in W$. The element $x'y'$, which belongs to A, is therefore contained in U. This shows that $xy \in \overline{A}$.

Next, if $x \in \overline{A}$ and U is an arbitrary neighborhood of x^{-1}, then there exists a neighborhood V of x such that $V^{-1} \subseteq U$. Further, there exists an element $x' \in A$ such that $x' \in V$, and therefore the element x'^{-1}, which belongs to A, is contained in U. This shows that $x^{-1} \in \overline{A}$. Thus, $\overline{A}$ proves to be a subgroup of G.

Now let A be normal in G. If $x \in \overline{A}$, $g \in G$, and U is an arbitrary neighborhood of $g^{-1}xg$, then there exists a neighborhood V of x such that $g^{-1}Vg \subseteq U$. In V we can find an element $x' \in A$, and therefore U contains the element $g^{-1}x'g \in g^{-1}Ag = A$. This shows that $g^{-1}xg \in \overline{A}$, i.e., that $\overline{A}$ is a normal subgroup.

Finally, let A be abelian. If $x, y \in \overline{A}$ and U is an arbitrary neighborhood of the element $x^{-1}y^{-1}xy$—the *commutator* of x and y—then it easily follows from the continuity of the operation in G that there exist neighborhoods V of x and W of y such that $V^{-1}W^{-1}VW \subseteq U$. Further, there exist elements $x', y' \in A$ such that $x' \in V$, $y' \in W$. Hence

$$1 = x'^{-1}y'^{-1}x'y' \in U,$$

that is, every neighborhood of $x^{-1}y^{-1}xy$ contains the unit element, and therefore $x^{-1}y^{-1}xy = 1$, so that $xy = yx$. //

Every open subgroup A of a topological group G is closed.

For let $x \in \overline{A}$. The coset Ax contains x and, by § 49.3, is open, so that it also contains an element of A. But this is only possible when $Ax = A$; hence $x \in A$, i.e., $\overline{A} = A$. //

7. A homomorphism φ of a topological group G onto a topological group G' is called *continuous* if it satisfies any of the following conditions, whose equivalence can be verified without difficulty:

1) For every subset A of G,

$$\overline{A}\varphi \subseteq \overline{A\varphi};$$

2) The complete inverse image $A'\varphi^{-1}$ of every closed subset A' of G' is closed in G;

3) The complete inverse image $A'\varphi^{-1}$ of every open subset A' of G' is open in G;

4) The complete inverse image $U'\varphi^{-1}$ of every neighborhood U' of the unit element of G' is open in G.

A continuous homomorphism φ of a topological group G onto a topological group G' is called *open* if it satisfies any of the following equivalent conditions:

1') The image $a\varphi$ of every open subset A of G is open in G';

2') The image $U\varphi$ of every neighborhood U of the unit element of G contains a neighborhood U' of the unit element of G'.

Finally, two topological groups G and G' are called *isomorphic* if there is a one-to-one correspondence between them that is an isomorphism for G and G' as groups (see § 10.1) and as topological spaces (see § 47.8).

8. Let A be a closed subgroup of a topological group G; M, the set of right cosets of A in G; φ, the mapping of G onto M that carries each element $x \in G$ into the coset Ax; Σ, a complete system of neighborhoods of G; and $\Sigma\varphi$, the system of images of the neighborhoods in Σ under the mapping φ.

The set M can be regarded as a topological space with $\Sigma\varphi$ as a complete system of neighborhoods.

To prove this, it is sufficient to employ the criterion of § 47.7. If $Ax \neq Ay$, then, since Ay is closed (see § 49.3), Σ contains a neighborhood U such that $x \in U$ and $U \cap Ay$ is empty. Then $U\varphi$ contains Ax, but does not contain Ay, so that condition α) is satisfied.

Further, let U and V be neighborhoods in Σ such that $U\varphi \cap V\varphi$ contains the coset Ax. Since, by § 49.3, the set aU is open for every $a \in A$, the product AU is also open; the same is true of AV. The element x is contained in $AU \cap AV$; consequently, we can find in this intersection a W of Σ containing x. Since $(AU)\varphi = U\varphi$ and $(AV)\varphi = V\varphi$, we have $W\varphi \subseteq (U\varphi \cap V\varphi)$, with $Ax \in W\varphi$. This shows that the condition β) is also satisfied. //

It is in the sense of this topology that we shall speak of the *space of right cosets* (or similarly, *of left cosets*) of a topological group with respect to a closed subgroup and, in particular, of the *topology in the factor group* of a topological group with respect to a closed normal subgroup.

❖ The factor group G/A of a topological group G with respect to a closed normal subgroup A is a topological group, and the natural homomorphism of G onto G/A is open.

If φ is an open homomorphism of a topological group G onto a topological group G', then the kernel A of this homomorphism is a closed normal subgroup and there exists an isomorphism ψ of G' onto the topological group G/A such that the product $\varphi\psi$ coincides with the natural homomorphism of G onto G/A.

The factor group G/A of a topological group G is discrete if and only if A is open. ❖

9. A topologization of a group G is called *linear* (or *totally linear*) if it can be given by a complete system of neighborhoods of the unit element consisting of subgroups (or normal subgroups) of the group.

The discrete topologization of every group is, obviously, totally linear.

A system Σ of normal subgroups of a group G is a complete system of neighborhoods of the unit element in some totally linear topologization of the group if and only if

1′) *The intersection of all the normal subgroups in Σ consists of the unit element of G only;*

2′) *The intersection of any two normal subgroups in Σ contains a third normal subgroup belonging to Σ.*

The necessity of the conditions 1′) and 2′) is obvious. On the other hand, if a system Σ of normal subgroups satisfies these conditions, then the

system Σ_0 of all the cosets with respect to all the normal subgroups in Σ satisfies the conditions α) and β) of § 47.7 and so is a complete system of neighborhoods for some topology in G. In fact, if $a, b \in G$ $(a \neq b)$, then by 1'), we can find a normal subgroup A in Σ such that $a^{-1}b \notin A$, and then a and b lie in distinct cosets of A. But if $a \in G$ and $A, B \in \Sigma$, then by 2'), Σ contains a normal subgroup C such that $C \subseteq A \cap B$; and then $aC \subseteq (aA \cap aB)$.

It remains to show that the group operations of G are continuous in this topology (see § 49.1). If $a, b \in G$ and $A \in \Sigma$, then $aA \cdot bA = (ab)A$, and this proves the continuity of multiplication; if $a \in G$ and $A \in \Sigma$, then $(aA)^{-1} = a^{-1}A$, and this implies the continuity of the operation of taking the inverse element. //

From this theorem it follows, in particular, that *every decreasing sequence of normal subgroups of a group G whose intersection is equal to E determines a totally linear topologization of the group.*

10. *Every linear topological group G is totally disconnected* (see § 48.6).

For let Σ be a complete system of neighborhoods of the unit element of G. If $a, b \in G$ $(a \neq b)$, then Σ contains a subgroup A such that $a^{-1}b \notin A$, i.e., $aA \neq bA$. By § 49.3, every coset xA—including, in particular, aA—is an open set. The set $G \setminus aA$, as the union of all left cosets of A in G other than aA, is also open. Thus, we have obtained a decomposition of G into two disjoint open (closed) subsets with $a \in aA$ and $b \in G \setminus aA$. //

❖ Every locally compact (see § 48.4) totally disconnected topologization of an arbitrary group is linear.

Every compact (see § 48.3) totally disconnected topologization of an arbitrary group is totally linear. ❖

§ 50. Connection Between Topology and Valuations in Rings and Fields

A ring R is called a *topological ring* if a topology is given in R under which the additive group of R is a topological group and the multiplicative groupoid of the ring is a topological groupoid (see § 49.1).

An associative field K is called a *topological field* if a topology is given in K under which it is a topological ring and if, in addition, the operation of taking the inverse of an element is continuous in the sense of § 49.1. The multiplicative group of a topological field is therefore a topological group. The definition of a topological field can easily be put into a form in which it is applicable to non-associative fields and to division rings as well.

The reader will have no difficulty in formulating the definition of an isomorphism of topological rings. More generally, many problems that were discussed in the preceding section for topological groups can also be raised for topological rings and fields. However, we shall not dwell on this here.

2. *Every ring R with a pseudo-valuation $w(a)$* (see § 44.7) *with real values is a topological ring, and a complete system of neighborhoods of the zero is constituted by the set U_α (where α is an arbitrary positive real number) of those elements a of R for which $w(a) < \alpha$.*

For we know from § 49.3 that if the additive group of a ring R is a topological group, then not only are the U_α open sets but so also are all the sets of the form $a + U_\alpha$. Let us show that the system of all these sets (for arbitrary $a \in R$ and arbitrary positive real α) is indeed a complete system of neighborhoods for a suitable topology in R (see § 47.7).

Let $a, b \in R$ $(a \neq b)$. If $w(b-a) = \gamma$, then $\gamma > 0$. The set $a + U_\gamma$ contains a, because $0 \in U_\gamma$, but does not contain b, because $b - a \notin U_\gamma$.

Next, let

$$a \in (b + U_\beta) \cap (c + U_\gamma).$$

Then $w(a-b) = \delta_1 < \beta$, $w(a-c) = \delta_2 < \gamma$. We set

$$\alpha = \min(\beta - \delta_1,\ \gamma - \delta_2).$$

If $x \in R$, $w(x) < \alpha$, then

$$a + x = b + (a-b) + x,$$

where

$$w[(a-b) + x] < \delta_1 + \alpha \leqq \delta_1 + (\beta - \delta_1) = \beta.$$

This shows that $a + U_\alpha \subseteq b + U_\beta$. Similarly, $a + U_\alpha \subseteq c + U_\gamma$.

Thus, we have obtained a topology in R, and we see that every neighborhood of the element $a \in R$ contains a neighborhood of the form $a + U_\alpha$. We shall now prove the continuity of the ring operations of R in this topology.

Let $a, b \in R$ and let $a + b + U_\gamma$ be a given neighborhood of $a + b$. If $\delta = \gamma/2$ and $x, y \in U_\delta$, then

$$w(x + y) \leqq w(x) + w(y) < 2\delta = \gamma.$$

This shows that

$$(a + U_\delta) + (b + U_\delta) \subseteq a + b + U_\gamma,$$

so that the continuity of addition is proved.

Next, let $a \in R$ and let $-a + U_\alpha$ be a given neighborhood of $-a$. If $x \in U_\alpha$, then since $w(-x) = w(x)$ and $-x \in U_\alpha$, we have

$$-(a+U_\alpha)\subseteq -a+U_\alpha.$$

This proves the continuity of the operation of taking the opposite element.

Finally, if $a, b\in R$, $w(a)=\alpha$, $w(b)=\beta$, and if $ab+U_\gamma$ is a given neighborhood of ab, then we choose δ as a positive real number for which

$$a\delta+\delta\beta+\delta^2<\gamma.$$

If $x, y\in U_\delta$, then

$$w(ay+xb+xy)\leqq w(a)w(y)+w(x)w(b)+w(x)w(y)<\gamma,$$

i.e.,

$$(a+U_\delta)(b+U_\delta)\subseteq ab+U_\gamma,$$

and this shows the continuity of the multiplication. //

3. *If $w(a)$ is a real valuation* (see § 44.1) *in an associative field K, then K is a topological field in the topology defined in* § 50.2.

We already know that K is a topological ring, so that it remains to show the continuity of the operation of taking the inverse element.

We note, first of all, that it follows from

$$w(a)=w(a\cdot 1)=w(a)\cdot w(1)\qquad (a\in K)$$

that $w(1)=1$. Hence, for $a\neq 0$, because

$$w(a)\cdot w(a^{-1})=w(aa^{-1})=w(1)=1,$$

we have

$$w(a^{-1})=[w(a)]^{-1}.$$

Suppose now that $a\in K$, $a\neq 0$, $w(a)=\alpha$, and that $a^{-1}+U_\gamma$ is a given neighborhood of a^{-1}. We set

$$\delta=a\gamma/(a^{-1}+\gamma);\tag{1}$$

clearly, $0<\delta<\alpha$. Let $w(x)<\delta$. Then $a+x\neq 0$, and there exists a $y\in K$ such that

$$(a+x)^{-1}=a^{-1}+y.$$

Hence

$$y=(a+x)^{-1}-a^{-1}=a^{-1}[a-(a+x)](a+x)^{-1}=$$
$$=a^{-1}(-x)(a+x)^{-1}.$$

However, $w(a^{-1})=\alpha^{-1}$, $w(-x)=w(x)<\delta$. Further,

$$w(a)=w[(a+x)-x]\leqq w(a+x)+w(x),$$

and hence

$$w(a+x) \geqq w(a) - w(x) > \alpha - \delta > 0,$$

so that

$$w[(a+x)^{-1}] = [w(a+x)]^{-1} < (\alpha - \delta)^{-1}.$$

Thus,

$$w(y) < \alpha^{-1}\delta(\alpha - \delta)^{-1},$$

which, after replacement of δ by its expression (1), yields

$$w(y) < \gamma.$$

Thus, we have shown that $(a + U_\delta)^{-1} \subseteq a^{-1} + U_\gamma$. //

4. The usual valuation of the field of real numbers and the field of complex numbers by means of the absolute value *induces* (in the sense of the theorems proved above) the usual topologization of these fields. The usual valuation of the quaternion field (see § 44.1) induces a topologization of this field that coincides on its additive group with the usual topologization of a four-dimensional euclidean space (see § 47.10). It is this topologization of these three fields that we have in mind when we speak of them as topological fields.

In the field of rational numbers and the field of p-adic numbers (see § 19.12) their p-adic valuation (see §§ 45.6 and 45.8) also induces a topologization. This topologization (as well as the topologization, generated by it, of the ring of p-adic integers and of the additive groups of these fields and rings) is called their *p-adic topologization*. Since the p-adic valuation is logarithmic and has integer values, a system of neighborhoods of zero in this topology is formed by the sets U_k (where k is an arbitrary integer) of those numbers a (rational or p-adic) whose p-adic value is strictly larger than k.

5. Let K be an associative topological field. An element $a \in K$ is called *topologically nilpotent* if the sequence of its powers a^n $(n = 1, 2, \ldots)$ converges to zero, that is, if every neighborhood U of zero contains all, except possibly a finite number, of these powers. An element $a \in K$ $(a \neq 0)$ is called *neutral* if neither a nor a^{-1} is topologically nilpotent. Finally, a set A $(A \subset K)$ is called *bounded on the right* if for every neighborhood U of zero we can find a neighborhood V of zero such that $AV \subseteq U$.

We shall prove the following theorem [I. Kaplansky, *Duke Math. J.*, vol. 14 (1947), pp. 527-41; for topological fields, in a somewhat different formulation: I. R. Shafarevich, *DAN SSSR*, vol. 40 (1943), pp. 149-51].

Let K be an associative topological field. The topology of this field is induced by a real valuation if and only if the following conditions hold:

1) *The set N of all topologically nilpotent elements of K is open and bounded on the right;*

2) *If the element a is topologically nilpotent and b is topologically nilpotent or neutral, then the product ba is topologically nilpotent.*

6. The *necessity* of these conditions can be verified without difficulty. If the topology of K is induced by the real valuation $w(a)$, then the fact that a is topologically nilpotent is equivalent to $w(a) < 1$, and that it is neutral, to $w(a) = 1$. Condition 2) follows easily from this. The fact that N is bounded on the right follows from the inclusion $NU_\gamma \subseteq U_\gamma$. Finally, N is open, because $N = U_1$.

7. We now proceed to the proof of the *sufficiency* of these conditions. To begin with, we show that if $a \in N$ and $a \neq 0$, then $a^{-1} \notin N$. Suppose that this is not the case, and let U be an arbitrary neighborhood of zero. Then, by virtue of the equation $0 \cdot 0 = 0$ and the continuity of multiplication, there exist neighborhoods of zero, V and W, such that $V \cdot W \subseteq U$. Since for sufficiently large natural numbers n we have $a^n \in V$, $a^{-n} \in W$, we see that $1 \in U$; but this is impossible, because U was an arbitrary neighborhood of zero.

We denote by M the set of all neutral elements. Condition 2) of the theorem can now be written in the form

$$(N \cup M)N \subseteq N. \tag{2}$$

Clearly, $1 \in M$. From the definition of a neutral element it follows that if $a \in M$, then $a^{-1} \in M$ as well. Next, let $a, b \in M$. If $ab \in N$, then by (2), $a^{-1}(ab) = b \in N$; but if $(ab)^{-1} \in N$, then $b(ab)^{-1} = a^{-1} \in N$. Thus, $ab \in M$, i.e.,

$$MM \subseteq M. \tag{3}$$

Now let $a \in N$, $b \in M$. If $ab \in M$, then by (3) and the fact that $b^{-1} \in M$, we have $(ab)b^{-1} = a \in M$. But if $(ab)^{-1} \in N$, then $b(ab)^{-1} = a^{-1} \in N$. Thus, $ab \in N$, i.e.,

$$NM \subseteq N. \tag{4}$$

From (2)-(4) it follows that

$$(N \cup M)(N \cup M) \subseteq (N \cup M). \tag{5}$$

Next, let $a \in N$ and let x be an arbitrary non-zero element of K. If U is a given neighborhood of zero, then it follows from $x^{-1} \cdot 0 \cdot x = 0$ and the continuity of multiplication that there exists a neighborhood of zero V such

that $x^{-1}Vx \subseteq U$. Since $a^n \in V$ for sufficiently large natural numbers n, we have $x^{-1}a^n x = (x^{-1}ax)^n \in U$; hence $x^{-1}ax \in N$, i.e.,

$$x^{-1}Nx \subseteq N. \tag{6}$$

But if $a \in M$, $x \neq 0$, then $x^{-1}ax \in M$, because by (6) it would follow from $b = x^{-1}ax \in N$ that $xbx^{-1} = a \in N$ and from $b^{-1} \in N$ that $a^{-1} \in N$. Thus,

$$x^{-1}Mx \subseteq M. \text{ //} \tag{7}$$

8. It follows from the above results, and in particular from (3) and (7), that *M is a normal subgroup of the multiplicative group K^* of the non-zero elements of K.* We can therefore consider the factor group K^*/M. We note that this factor group is trivial if and only if N consists of the zero only, that is—in virtue of condition 1)—if the topology of K is discrete: we have to bear in mind that K cannot contain *nilpotent* elements, that is, elements a such that $a \neq 0$, but $a^n = 0$ for some natural number n.

Because of (5), the set $N \cup M$ (without the zero) splits into complete cosets of M, and these cosets form a sub-semigroup in the factor group K^*/M. With the help of (6) and (7) it is easy to verify the conditions 1)-4) of § 41.3, so that this sub-semigroup is the semigroup of positive elements for a suitable linear ordering of K^*/M.

This ordering turns out to be archimedean (see § 43.3). For if $a, b \in K^*$ and the cosets aM and bM are strictly positive, then $a \in N$, $b \in N$. Let U be a given neighborhood of zero; we may assume that $U \subseteq N$, because N is open. In virtue of $0 \cdot b^{-1} = 0$ there exists a neighborhood of zero V such that $Vb^{-1} \subseteq U$, and since it follows from $a \in N$ that there exists a natural number k such that $a^k \in V$, we have $a^k b^{-1} \in U$ and therefore $a^k b^{-1} \in N$. Hence in the ordered group K^*/M we have $(aM)^k > bM$, and the group is archimedean. //

9. On account of Hölder's Theorem (§ 43.4), the group K^*/M is isomorphic to a subgroup of the additive group of real numbers in its natural ordering. This determines a homomorphic mapping φ of K^* into the additive group of real numbers under which the elements of M, and these elements only, are mapped into zero and the images of the non-zero elements of N (and of these elements only) are strictly positive. For the time being, let λ be an arbitrary positive real number. For $a \in K$ we set

$$\left.\begin{aligned} w(a) &= 2^{-\lambda \cdot a\varphi} \qquad (a \neq 0), \\ w(0) &= 0. \end{aligned}\right\} \tag{8}$$

It is clear that the conditions 1) and 2) of § 44.1 are satisfied. Later we shall show that we can choose λ in such a way that the condition 3) is also satisfied.

It follows easily from (8) that $w(a)=1$ if and only if $a\in M$, and $w(a)<1$ if and only if $a\in N$. As in § 50.2, we denote by U_α (where α is a positive real number) the set of those elements $a\in K$ for which $w(a)<\alpha$. If the topology of K is discrete, then for $\alpha<1$ we have $U_\alpha=0$.

Suppose, then, that the topology of K is not discrete. Consider an arbitrary U_α. In K there exists an element b such that $w(b)>\alpha^{-1}$. Since N is an open set, we can find a neighborhood of zero U such that $bU\subseteqq N=U_1$. Therefore, $U\subseteqq U_\alpha$.

Conversely, let U be an arbitrary neighborhood of zero in K. Since N is bounded on the right and we have assumed that the topology in K is non-discrete, there exists a $c\neq 0$ such that $Nc\subseteqq U$. Hence, if $\alpha=w(c)$, then $U_\alpha\subseteqq U$, because $U_\alpha=U_1c=Nc$.

Thus, *if λ is chosen so that the valuation becomes $w(a)$, then the topology induced by this valuation in the sense of* § 50.2 *is the same as the original topology of K* (see the definition of equivalent systems of neighborhoods in § 47.8).

10. Continuing to regard λ as chosen arbitrarily, we shall now prove the existence of a number l such that for all $a\in K$ we have the inequality

$$w(1+a)\leqq l(1+w(a)). \tag{9}$$

For if such an l does not exist, then we could find a sequence of elements $a_n\in K$, $a_n\neq -1$ $(n=1, 2, \ldots)$, such that, as $n\to\infty$,

$$[1+w(a_n)]/[w(1+a_n)]=1/[w(1+a_n)]+w(a_n)/[w(1+a_n)]\to 0,$$

i.e.,

$$1/[w(1+a_n)]=w((1+a_n)^{-1})\to 0,$$
$$w(a_n)/[w(1+a_n)]=w(a_n(1+a_n)^{-1})\to 0.$$

Hence it would follow, by what we have proved in the preceding subsection, that the sequences

$$(1+a_n)^{-1} \quad\text{and}\quad a_n(1+a_n)^{-1} \qquad (n=1, 2, \ldots) \tag{10}$$

converge to zero in the topology of K; but this is impossible, because addition in K is continuous and the sum of the n-th terms of the sequences (10) is equal to 1.

It follows from (9) that if at least one of the elements $a, b\in K$ is different from zero, then

$$w(a+b) \leqq l(w(a)+w(b)). \tag{11}$$

For if $a \neq 0$, then

$$w(a+b) = w[a(1+a^{-1}b)] \leqq lw(a)[1+w^{-1}(a)w(b)] = l(w(a)+w(b)).$$

It is clear that (11) is also true for $a = b = 0$. It follows from (11) that

$$w(a+b) \leqq 2l \max(w(a), w(b)). \tag{12}$$

11. There exists a positive real number ν sufficiently small that $(2l)^\nu \leqq 2$. In the definition (8) we replace λ by $\lambda\nu$. The number $w(a)$ is then raised to the power ν; if from now on we refer the symbol w to the new choice of λ, we obtain from (12)

$$w(a+b) \leqq 2 \max(w(a), w(b)). \tag{13}$$

From this it follows easily that for $k = 1, 2, \ldots,$

$$w(a_1 + a_2 + \ldots + a_{2^k}) \leqq 2^k \max_{1 \leqq i \leqq 2^k} w(a_i). \tag{14}$$

We shall use the symbol n to denote the n-fold of the unit element of K; this element will only occur under the functional symbol w, so that there is no danger of confusing it with the natural number n. We also note that K is not assumed to be a field without characteristic. From (14) it follows that

$$w(2^k) \leqq 2^k \qquad (k = 1, 2, \ldots). \tag{15}$$

Let us show that for every natural number n

$$w(n) \leqq 2n. \tag{16}$$

For, let $2^k \leqq n < 2^{k+1}$. We shall prove (16) by induction on k. For $k = 0$, the statement is true. Since $n = (n - 2^k) + 2^k$, we derive from (13) and (15), provided $w(n-2^k) \leqq w(2^k)$, that

$$w(n) \leqq 2w(2^k) \leqq 2 \cdot 2^k \leqq 2n.$$

But if $w(n-2^k) \geqq w(2^k)$, then we can apply the induction hypothesis, inasmuch as $n - 2^k < 2^k$; and so we see that

$$w(n) \leqq 2w(n-2^k) \leqq 2 \cdot 2(n-2^k) = 4n - 2^{k+2} < 2n.$$

12. We now choose an arbitrary element $a \in K$ and apply (14) to the expansion of the element $(1+a)^{n-1}$ by the binomial formula, making the assumption that n is a power of 2. The binomial formula is applicable to

the case in question, because the elements a and 1 are permutable. Since, by (16),

$$w\left(\left(\begin{matrix} n-1 \\ k \end{matrix}\right)\right) \leqq 2\left(\begin{matrix} n-1 \\ k \end{matrix}\right),$$

where $\left(\begin{matrix} n-1 \\ k \end{matrix}\right)$ is the binomial coefficient, we have

$$w((1+a)^{n-1}) \leqq n \max w\left(\left(\begin{matrix} n-1 \\ k \end{matrix}\right)a^k\right) = n \max \left[w\left(\left(\begin{matrix} n-1 \\ k \end{matrix}\right)\right)w(a^k)\right] \leqq$$

$$\leqq 2n \max\left(\left(\begin{matrix} n-1 \\ k \end{matrix}\right)w^k(a)\right) \leqq 2n(1+w(a))^{n-1}.$$

Hence it follows that

$$[w(1+a)]^{n-1} \leqq 2n(1+w(a))^{n-1},$$

and since n can be as large a power of 2 as we please,

$$w(1+a) \leqq 1 + w(a).$$

In our present case of a real norm, the inequality 3_2) of § 44.1, is equivalent to the condition 3) in the definition of a valuation: for we have $w(-1) = 1$, so that the equation 3_1) of § 44.1 is satisfied. We can now derive 3_2) in the same way as (11) was derived from (9). This completes the proof of the theorem of § 50.5. //

❖ If an associative topological field is locally compact (see § 48.4), then its topology is induced by a real valuation [I. Kaplansky, *Duke Math. J.*, vol. 14 (1947), pp. 527-41].

The field of real numbers, the field of complex numbers, and the quaternion field, considered in their natural topologies (see § 50.4), are the only connected (see § 48.6) locally compact associative topological fields [L. S. Pontrjagin, *Ann. of Math.*, vol. 33 (1932), pp. 163-74].

Let F be a topological field. The topology of F is induced by a non-archimedean valuation (see § 44.3) with values in a linearly ordered group supplemented by a zero if and only if: 1) F contains a neighborhood of zero generated by a bounded additive subgroup; 2) If a subset $A \subset F$ does not intersect some neighborhood of zero, then A^{-1} is bounded [D. Zelinski, *Bull. Amer. Math. Soc.*, vol. 54 (1948), pp. 1145-50]. ❖

§ 51. Galois Connections. The Fundamental Theorem of Galois Theory

1. We say that a *Galois connection* exists between two partially ordered sets M and M' if there are mappings $\varphi: M \to M'$ and $\psi: M' \to M$ satisfying (for arbitrary $a, b \in M$ and $a', b' \in M'$) the following conditions:

1) If $a \leqq b$, then $a\varphi \geqq b\varphi$;
   If $a' \leqq b'$, then $a'\psi \geqq b'\psi$;
2) $a\varphi\psi \geqq a$, $a'\psi\varphi \geqq a'$.

This concept is closely connected with that of a closure in a partially ordered set (see § 47.1):

If a Galois connection exists between partially ordered sets M and M', then the equations

$$\begin{aligned} \bar{a} &= a\varphi\psi \qquad (a \in M), \\ \bar{a}' &= a'\psi\varphi \qquad (a' \in M') \end{aligned} \tag{1}$$

determine closure relations in M and M', respectively. If M_0 and M_0' are the systems of all elements of M and M', respectively, that are closed under these closures, then φ is an inverse isomorphism of M_0 onto M_0', and ψ an inverse isomorphism of M_0' onto M_0, and on these sets the mappings φ and ψ are inverse to each other.

Let us show that the equations (1) introduce closure relations into M and M'. It is clear that 2) immediately leads to the condition 1_0) of § 47.1, and that 1) implies the condition 2_0). On the other hand, it follows from 2), since $a\varphi \in M'$, that

$$a\varphi\psi\varphi \geqq a\varphi,$$

and since, by 1), it follows from $a\varphi\psi \geqq a$ that

$$a\varphi\psi\varphi \leqq a\varphi,$$

we have in fact

$$a\varphi\psi\varphi = a\varphi, \tag{2}$$

and therefore

$$a\varphi\psi\varphi\psi = a\varphi\psi \qquad (a \in M). \tag{3}$$

This shows that the condition 3_0) of § 47.1 is also satisfied.

The systems M_0 and M_0' of elements that are closed in M and M' relative to the closures constructed above consist, respectively, of the elements $a\varphi\psi$, where $a \in M$, and $a'\psi\varphi$, where $a' \in M'$. The equation (2), rewritten in the form

$$(a\varphi)\psi\varphi = a\varphi,$$

shows that the image under φ of an arbitrary element a of M_0 is closed in M'. Hence φ maps M_0 into M_0'. Similarly, the mapping ψ carries M_0' into M_0.

But from (3) and the analogous equation

$$a'\psi\varphi\psi\varphi = a'\psi\varphi \qquad (a' \in M'),$$

it follows that φ and ψ, considered only on M_0 and M_0', are in fact one-to-one mappings inverse to each other. They are even inverse isomorphisms, as follows from 1). //

2. We shall give here only one of the numerous applications that Galois connections have in algebra. We consider a commutative field K with a distinguished subfield P and we denote by $G(K,P)$ the set of all automorphisms of K that leave every element of P invariant. This set is obviously a group under the multiplication of automorphisms. The group $G(K,P)$ is called the *Galois group* of K over P.

We denote by M the set of all subsets A of K that contain P,

$$P \subseteqq A \subseteqq K,$$

and by M' the set of all subsets A' of the group $G(K,P)$, where both these sets are considered in their set-theoretical partial ordering. For every $A \in M$ we denote by $A\varphi$ the collection of those automorphisms of $G(K,P)$ that leave the set A elementwise fixed. On the other hand, for every $A' \in M'$ we denote by $A'\psi$ the set of all elements of K that are left fixed by all the automorphisms of A'.

It is clear that *$A\varphi$ is a subgroup of the Galois group $G(K,P)$ and $A'\psi$ a subfield of K containing P.* It is also readily verified that the mappings φ and ψ satisfy the conditions 1) and 2) of § 51.1. Therefore, we have established a Galois connection between M and M', which determines closures in each of these two sets. Finally, *the mappings φ and ψ are isomorphisms inverse to each other* between the systems $M_0 \subset M$ and $M_0' \subset M'$ of closed elements, that is, *between a certain system of subfields of K containing P and a certain system of subgroups of the Galois group $G(K,P)$.*

3. Of particular interest is the case where the subfields intermediate between P and K as well as the subgroups of the Galois group all prove to be closed; in that case, the survey of all subfields of K containing P reduces to the survey of all subgroups of $G(K,P)$. Let us examine one such case.

We shall only discuss commutative fields without characteristic (see § 18.11). We shall also assume that the reader is familiar, from a course in higher algebra, with the elements of the algebra of polynomials over

commutative fields [see, for example, A. G. Kurosh, *A Course of Higher Algebra* (6th ed. [in Russian], 1959; English translation in prep.), §§ 47-49; this book will be quoted as 'H. A.'].

As above, let $P \subset K$. An element $\alpha \in K$ is called *algebraic* over P if it is the root of a polynomial in the ring $P[x]$ and therefore the root of a uniquely determined irreducible polynomial $\varphi(x)$ with leading coefficient 1. The degree of this polynomial is called the *degree* of α, and its roots in K or in some *extension field* of K are called the elements *conjugate* to α.

The field K is called an *algebraic extension* of P if every element of K is algebraic over P. K is a *finite extension* of P—more precisely, *an extension of degree* n—if K is a finite-dimensional linear algebra over P (see § 46.6) of dimension n. The degree of K over P is denoted by the symbol $(K : P)$.

Every finite extension is algebraic.

For if $(K : P) = n$, then any $n + 1$ elements of K are linearly dependent over P. In particular, if $\alpha \in K$, then the elements $1, \alpha, \alpha^2, \ldots, \alpha^n$ are linearly dependent over P, and this means that α is a root of a polynomial of degree n in $P[x]$.

4. THE THEOREM OF THE PRIMITIVE ELEMENT: *If a commutative field K without characteristic is an algebraic extension of P and is generated by the adjunction of a finite number of elements to P, then K can be generated by the adjunction of a single element to P.*

It is sufficient to examine the case of the adjunction of two elements. Let

$$K = P(\alpha, \beta).$$

If α and β, respectively, are roots of the polynomials $\varphi(x)$ and $\psi(x)$ irreducible over P, then there exists an extension L of K (for example, the splitting field of the product $\varphi(x)\psi(x)$; see, for example, H. A., § 49) in which both these polynomials have all their roots.

Let

$$\alpha_1 = \alpha, \ \alpha_2, \ \ldots, \ \alpha_k$$

be all the roots of $\varphi(x)$ in L; they are distinct, because $\varphi(x)$ has no multiple roots. Similarly

$$\beta_1 = \beta, \ \beta_2, \ \ldots, \ \beta_l$$

are all the roots of $\psi(x)$ in L. Since P, as a field without characteristic, is infinite, we can find an element $c \in P$ such that

$$\alpha_i + c\beta_j \neq \alpha + c\beta \tag{4}$$

if at least one of the indices i, j is different from 1.

We set

$$\gamma = \alpha + c\beta. \tag{5}$$

It is clear that $P(\gamma) \subseteq K$. We now consider the polynomial

$$\overline{\varphi}(x) = \varphi(\gamma - cx)$$

with coefficients in $P(\gamma)$. Since, by (5),

$$\overline{\varphi}(\beta) = \varphi(\gamma - c\beta) = \varphi(\alpha) = 0,$$

and since we know that $\psi(\beta) = 0$, the polynomials $\overline{\varphi}(x)$ and $\psi(x)$ have a common root β. This is their only common root: if $\overline{\varphi}(\beta_j) = 0$ $(j \neq 1)$, then

$$\varphi(\gamma - c\beta_j) = 0,$$

that is, $\gamma - c\beta_j$ is equal to some α_i, in contradiction to (4).

Hence it follows that $x - \beta$ is the greatest common divisor of the polynomials $\overline{\varphi}(x)$ and $\psi(x)$. The coefficients of these polynomials, and hence of their greatest common divisor, lie in the field $P(\gamma)$, so that $\beta \in P(\gamma)$. Hence it follows, by (5), that $\alpha \in P(\gamma)$ as well. Thus,

$$K = P(\alpha, \beta) = P(\gamma),$$

as required. //

5. An algebraic extension K of a commutative field P is called a *normal extension* if every irreducible polynomial in $P[x]$ having at least one root in K has all its roots in K, that is, splits into linear factors over K.

❖ A commutative field K is a normal extension of its subfield P if and only if K is generated by the adjunction to P of all the roots of a set of polynomials in $P[x]$. ❖

6. *If a commutative field K without characteristic is a normal extension of its subfield P, of degree n, then the Galois group $G(K, P)$ is finite and of order n.* For, by § 51.3, K is algebraic over P, and by § 51.4, it has a primitive element

$$K = P(\alpha).$$

Since $(K : P) = n$, every element $\beta \in K$ has a unique expression of the form

$$\beta = b_0 + b_1\alpha + b_2\alpha^2 + \ldots + b_{n-1}\alpha^{n-1} \quad (b_0, b_1, b_2, \ldots, b_{n-1} \in P) \tag{6}$$

and therefore α is of degree n over P. Since K is normal over P and since

it is a field without characteristic, K contains n distinct elements conjugate to α:

$$\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n. \tag{7}$$

Clearly,

$$K = P(\alpha_i) \qquad (i = 1, 2, \ldots, n).$$

As is well known (see H. A., § 49), there exists a uniquely determined automorphism of K that leaves the elements of P invariant and carries α into α_i $(i = 1, 2, \ldots, n)$, namely the automorphism φ_i that maps the element β of (6) into

$$\beta_i = b_0 + b_1\alpha_i + b_2\alpha_i^2 + \ldots + b_{n-1}\alpha_i^{n-1}.$$

This yields n elements of the Galois group $G(K, P)$. On the other hand, let φ be an arbitrary automorphism of K belonging to $G(K, P)$, and let $\alpha\varphi = \alpha'$. If

$$f(x) = x^n + a_1x^{n-1} + \ldots + a_n$$

is the irreducible polynomial in $P[x]$ which has α as a root, then since $a\varphi = a$ for all $a \in P$,

$$f(\alpha') = \alpha'^n + a_1\alpha'^{n-1} + \ldots + a_n =$$
$$= (\alpha^n + a_1\alpha^{n-1} + \ldots + a_n)\varphi = [f(\alpha)]\varphi = 0\varphi = 0.$$

The element α' therefore coincides with one of the elements α_i of (7), and hence so does the automorphism φ with the corresponding automorphism φ_i. //

7. It is clear that the fields Q intermediate between P and K,

$$P \subseteqq Q \subseteqq K, \tag{8}$$

form a lattice under set-theoretical inclusion. Also, it clearly follows from $K = P(\alpha)$ that $K = Q(\alpha)$ for every Q satisfying (8).

If the field K is finite over P, then an intermediate field Q is finite over P, K is finite over Q, and

$$(K : P) = (K : Q)(Q : P).$$

That Q is finite over P follows from the fact that Q is a subspace of the vector space K over P. That K is finite over Q follows from the fact that we obtain a basis of K over Q by choosing, in a basis K over P, a maximal subsystem that is linearly independent over Q. Finally, let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be a basis of K over Q, and $\beta_1, \beta_2, \ldots, \beta_l$ a basis of Q over P. Then for every $\gamma \in K$ we have

$$\gamma = \sum_{i=1}^{k} b_i \alpha_i \qquad (b_i \in Q;\ i = 1, 2, \ldots, k), \tag{9}$$

and since

$$b_i = \sum_{j=1}^{l} a_{ij} \beta_j \qquad (a_{ij} \in P), \tag{10}$$

we have

$$\gamma = \sum_{i=1}^{k} \sum_{j=1}^{l} a_{ij} (\alpha_i \beta_j). \tag{11}$$

The system of kl elements $\alpha_i \beta_j$ is linearly independent over P. This can be seen as follows: If the right-hand side of (11) is equal to zero, then the linear independence of the system $\alpha_1, \alpha_2, \ldots, \alpha_k$ over Q shows that the coefficients b_i in (9) are equal to zero, where b_i has the form (10); and then, since the system $\beta_1, \beta_2, \ldots, \beta_l$ is linearly independent over P, all the coefficients a_{ij} are equal to zero. //

If K is a normal field over P, then it is also normal over every intermediate field Q.

For let α be a root in K of the polynomial $g(x)$, irreducible over Q. The polynomial $f(x)$, irreducible over P, which has α as a root has all its roots in K, and since $g(x)$ is a divisor of $f(x)$ in $Q[x]$ (see H. A., § 48), $g(x)$ also splits into linear factors over K. //

8. THE FUNDAMENTAL THEOREM OF GALOIS THEORY: *If a field K without characteristic is a finite and normal extension of a field P, then the correspondence that assigns to every intermediate field Q (see (8)) the Galois group $G(K, Q)$ is an inverse isomorphism between the lattice of all intermediate fields and the lattice of all subgroups of a Galois group $G(K, P)$.*

For the mapping φ in the Galois connection considered in § 51.2 assigns to every intermediate field Q the subgroup of $G(K, P)$ that is the Galois group of K over Q:

$$Q\varphi = G(K, Q).$$

The fundamental theorem will therefore be proved if we can show that under the closures defined by this Galois connection all the intermediate fields as well as the subgroups of the Galois group are closed.

If Q is an arbitrary intermediate field, then

$$\begin{aligned} Q\varphi &= G(K, Q) = U \subseteq G(K, P), \\ U\psi &= Q' = Q\varphi\psi, \\ Q'\varphi &= G(K, Q') = U' = U\psi\varphi \subseteq G(K, P). \end{aligned}$$

Therefore $Q \subseteqq Q'$, $U \subseteqq U'$. We denote by s and s', respectively, the orders of the subgroups U and U'. Then

$$s \leqq s'$$

and, by § 51.7,

$$(K : Q) \geqq (K : Q'). \tag{12}$$

However, by § 51.6,

$$(K : Q) = s, \qquad (K : Q') = s'.$$

Therefore equality holds in (12), and hence $Q' = Q$.

Now let U be an arbitrary subgroup of $G(K, P)$. We set

$$\begin{aligned} U\psi &= Q, \\ Q\varphi &= U\psi\varphi = U' \subseteqq G(K, P). \end{aligned} \tag{13}$$

Then $U \subseteqq U'$, so that, as above, $s \leqq s'$. However, let $K = Q(\alpha)$. If

$$\varepsilon = \sigma_1, \sigma_2, \ldots, \sigma_s \tag{14}$$

are all the automorphisms, including the identity automorphism ε, in U, then α is a root of the polynomial of degree s

$$f(x) = (x - \alpha\sigma_1)(x - \alpha\sigma_2) \ldots (x - \alpha\sigma_s).$$

The coefficients of this polynomial expressed by Vieta's formulas do not change under any of the automorphisms (14): the multiplication of all the elements of U on the right by one of them only permutes these elements. Therefore, by the definition of the mapping and by (13),

$$f(x) \in Q[x],$$

that is, the degree of α over Q does not exceed s, so that

$$s' = (K : Q) \leqq s.$$

Thus, $s' = s$, i.e., $U' = U$. //

From this theorem it follows, in particular, that *under our conditions, the number of fields intermediate between P and K is finite.*

❖ Under the conditions of the fundamental theorem, an intermediate field Q is normal over P if and only if the corresponding subgroup $U = G(K, Q)$ is a normal subgroup of $G(K, P)$. The group $G(Q, P)$ is then isomorphic to the factor group

$$G(K, P)/U.$$

For fields without characteristic, the finiteness and normality of K over P are not only sufficient, but also necessary for the fundamental theorem to hold.

If a field K without characteristic is an algebraic and normal (but not necessarily finite) extension of P, then the correspondence

$$Q \to G(K, Q)$$

is an inverse isomorphism between the lattice of all subfields intermediate between P and K and the lattice of all subgroups of $G(K, P)$ that are closed in the linear topology of this group (see §49.9) defined as follows: a complete system of neighborhoods of the unit element is formed by the subgroups $G(K, Q_0)$, where Q_0 ranges over all the subfields of K that are finite normal extensions of P [W. Krull, *Math. Ann.*, vol. 100 (1928), pp. 687-98]. ❖

NOTE ON TERMINOLOGY

The present trend in a systematic mathematical terminology is to give the shortest name to the most general object under discussion and to add qualifying (positive) attributes to denote the more specialized objects (example: set—ordered set—well-ordered set). This is often in conflict with the historical development: when rings and fields first made their appearance in mathematics, the associative and commutative laws were part of their definition. At a later stage, wider objects of study had then to be denoted by (negative) attributes such as *non-associative* rings or *non-commutative* fields. For a number of years the term *skewfield,* or *sfield,* was used to refer sometimes to not necessarily commutative fields, sometimes to definitely non-commutative fields (and to add to the confusion, sometimes *division ring* was treated as a synonym). The Russian mathematical language is fortunate in having two distinct words: *тело* for a (general) field and *поле* for a commutative field. In this translation, too, I use *field* without implying the commutative law. But even then, both in Russian and English, the adjective *non-associative* refers to a wider, not a narrower, class of objects, and there are other similar compromises. The reader should consult, in particular, §§ 8.10 and 12.1. There are, in any case, limitations to the usefulness of a pedantically systematic scheme: no one has (yet) proposed to replace the term *groupoid* by *group* and to describe what we now know as a group as an 'associative group with division.'

K. A. H.

BIBLIOGRAPHY

[⁘ indicates translation into a Western European language (usually English); UMN = Usp. Mat. Nauk]

ALBERT, A. A. *Modern Higher Algebra.* Chicago: Univ. of Chicago Press, 1937.

ALBERT, A. A. *Structure of Algebras.* New York: American Math. Soc., 1939.

ALBERT, A. A. *Fundamental Concepts of Higher Algebra.* Chicago: Univ. of Chicago Press, 1956.

ALEKSANDROV, P. S. *Introduction to the Theory of Groups* [in Russian]. Moscow, 1938; 2nd edition, Moscow, 1951. Rumanian translation, Bucharest, 1954. German translation, Berlin, 1954. Ukrainian translation, Kiev, 1955. Polish translation, Warsaw, 1956. ⁘ English translation: ALEXANDROFF, P. S. *An Introduction to the Theory of Groups* [translated from the German]. London, 1959.

ALMEIDA COSTA A. *Abelian Groups, Noncommutative Rings and Ideals, Hypercomplex Systems and Representation.* Lisbon, vol. 1, 1942; vol. 2, 1948.

ALMEIDA COSTA, A. *Anéis associativos não comutativos* [vol. 3 of above]. Lisbon, 1955.

ANDREE, R. V. *Selections from Modern Abstract Algebra.* New York: Holt, Rinehart and Winston, 1958.

ARTIN, E. *Galois Theory.* Notre Dame Mathematical Lectures, No. 2, Notre Dame, 1942. 2nd edition, 1946. Chinese translation, Shangai, 1948.

ARTIN, E., NESBITT, C. J., and THRALL, R. M. *Rings with Minimum Condition.* Ann Arbor: Univ. of Michigan Press, 1944.

ARTIN, E. *Geometric Algebra.* New York: Interscience, 1957.

AZUMAYA, G. *Algebraic Theory of Prime Rings* [in Japanese]. Tokyo, 1951.

BAER, R. *Linear Algebra and Projective Geometry.* New York: Academic Press, 1952. Russian translation, Moscow, 1955.

BARBILIAN, D. *Teoria aritmetica a idealelor (în inele necomutative).* Bucharest, 1956.

BAUMGARTNER, L. *Gruppentheorie.* 3rd edition, Berlin: Walter de Gruyter, 1958.

BIRKHOFF, G. *Lattice Theory.* New York: American Math. Soc., 1940. 2nd edition, 1948. Russian translation, Moscow, 1952.

BIRKHOFF, G. and MACLANE, S. *A Survey of Modern Algebra.* New York: Macmillan, 1941. 2nd edition, 1953.

BOERNER, H. *Darstellungen von Gruppen.* Berlin-Göttingen-Heidelberg: Springer-Verlag, 1955. ⁘ English translation: *Representations of Groups.* Amsterdam: North-Holland, 1963.

BOREVICH, Z. I., and FADDEEV, B. K. *Homology Theory in Groups, I* [in Russian]. Vestnik Leningr. Univ., 1956, No. 7, 3-39.

BORŮVKA, O. *Uvod do theorie grup.* Prague, 1944. 2nd edition, 1952.

BORŮVKA, O. *Grundlagen der Gruppoid- und Gruppentheorie.* Berlin: Deutscher Verlag der Wissenschaften, 1960.

BOURBAKI, N. *Éléments de mathematique, Algèbre,* Chaps. 1-8. Paris: Hermann, 1942-58.

BRUCK, R. H. *A Survey of Binary Systems.* Berlin-Göttingen-Heidelberg: Springer, 1958.

BURNSIDE, W. *Theory of Groups.* 2nd edition, Cambridge, 1911. Repr., New York: Dover, 1955.

CARMICHAEL, R. D. *Introduction to the Theory of Groups of Finite Order.* Boston, 1937. Reprint, New York: Dover, 1956.

CARTAN, H. and EILENBERG, S. *Homological Algebra.* Princeton: Princeton Univ. Press, 1956. Russian translation, Moscow, 1960.

CHEN KUO-TSAI. *Foundations of Modern Algebra* [in Chinese]. Shanghai, 1952.

CHÂTELET, A. *Arithmétique et algèbre modernes.* Vol. 1 (*Notions fondamentales—Groupes*), Paris, 1954. Vol. 2 (*Anneaux et corps, calcul algébrique, idéaux et divisibilité*), Paris, 1956.

CHEBOTAREV, N. G. *Foundations of Galois Theory, part* 1 [in Russian]. Leningrad-Moscow, 1934. German translation, Groningen, 1950.

CHEBOTAREV, N. G. *Galois Theory* [in Russian]. Moscow-Leningrad, 1936.

CHEBOTAREV, N. G. *Introduction to the Theory of Algebras* [in Russian]. Moscow-Leningrad, 1949. Chinese translation, Peking, 1954.

CHERNIKOV, S. N. "Finiteness Conditions in the General Theory of Groups" [in Russian]. UMN 14:5 (1959), 45-96.

CHEVALLEY, C. *Théorie des groupes de Lie,* vol. 3, *Théorèmes généraux sur les algèbres de Lie.* Paris: Hermann, 1955. Russian translation, Moscow, 1958.

CHEVALLEY, C. *Fundamental Concepts of Algebra.* New York: Academic Press, 1957.

CLIFFORD, A. H. and PRESTON, G. G. *The Algebraic Theory of Semigroups,* vol. 1. Providence: Amer. Math. Soc., 1962.

COXETER, H. S. M. and MOSER, W. O. J. *Generators and Relations for Discrete Groups.* Berlin-Göttingen-Heidelberg: Springer-Verlag, 1957.

DEURING, M. *Algebren.* Berlin, 1935. Reprint, New York: Chelsea, 1948.

DICKSON, L. E. *Linear Groups.* Leipzig, 1901. Reprint, New York: Dover, 1958.

DUBREIL, P. *Algèbre,* vol. I. Paris: Gauthier-Villars, 1946. 2nd edition, 1954.

DUBREIL, P., JACOTIN, M. L., LESIEUR, L., and CROISOT, R. *Leçons sur la théorie des treillis, des structures algébriques ordonnées et des treillis géométriques.* Paris, 1953.

FUCHS, L. *Abelian Groups* [in English]. Budapest, 1958.

GLIVENKO, V. I. *Théorie générale des structures.* Paris, 1938.

GLUSHKOV, V. M. "The Structure of Locally Compact Groups and Hilbert's Fifth Problem" [in Russian]. UMN 12:2 (1957), 3-41. ❖ English translation in *Translations, Series II,* vol. 15. Providence: American Math. Soc., 1961.

GRAEV, M. I. "Theory of Topological Groups, I" [in Russian]. UMN 5:2 (1950), 3-56.

HALL, M. *Projective Planes and Related Topics.* Pasadena: Calif. Inst. of Technol., 1954.

HALL, M. *The Theory of Groups.* New York: Macmillan, 1959.

HAUPT, O. *Einführung in die Algebra.* 2nd edition, Leipzig, 1952.

HERMES, H. *Einführung in die Verbandstheorie.* Berlin-Göttingen-Heidelberg: Springer-Verlag, 1955.

JACOBSON, N. *The Theory of Rings.* New York, 1943. Russian translation, Moscow, 1947.

JACOBSON, N. *Lectures in Abstract Algebra,* 3 vols. Princeton: Van Nostrand, 1951, 1952, and [1964]. Chinese translation, Peking, 1960.

JACOBSON, N. *Structure of Rings.* New York: American Math. Soc., 1956. Russian translation, Moscow, 1961.

JACOBSON, N. *Lie Algebras.* New York: Interscience, 1962.

JOHNSON, R. E. *First Course in Abstract Algebra.* New York: Prentice-Hall, 1953.

JORDAN, C. *Traité des substitutions et des équations algébriques.* Paris, 1870. Reprint, Paris, 1957.

KAPLANSKY, I. *Infinite Abelian Groups.* Ann Arbor: Univ. of Michigan Press, 1954.

KAPLANSKY, I. *An Introduction to Differential Algebra.* Paris: Hermann, 1957. Russian translation, Moscow, 1959.

KOCHENDÖRFFER, R. *Einführung in die Algebra.* Berlin, 1955. 2nd edition, 1962.

KRULL, W. *Idealtheorie.* Berlin, 1935. Reprint, New York: Chelsea, 1948.

KUROSH, A. G. "The Theory of Infinite Groups, Its Development, and Some Recurrent Problems" [in Russian]. UMN 3 (1937), 5-15.

KUROSH, A. G. *Theory of Groups* [in Russian]. Moscow-Leningrad, 1944. 2nd edition, Moscow, 1953. German translation, Berlin, 1953. Hungarian translation, Budapest, 1955. ✣ English translation: *The Theory of Groups,* 2 vols. New York: Chelsea, 1955-1956, 2nd English edition, 1959-1960. Rumanian translation, Bucharest, 1959. Japanese translation, 2 vols., Tokyo, 1960-1961.

KUROSH, A. G. "The Present State of the Theory of Rings and Algebras" [in Russian]. UMN 6:2 (1951), 3-15.

KUROSH, A. G. and CHERNIKOV, S. N. "Solvable and Nilpotent Groups" [in Russian]. UMN 2:3 (1947), 18-59. ✣ English translation reprinted in *Translations, Series I,* vol. 1: *Algebra.* Providence: American Math. Soc., 1962.

LEDERMANN, W. *Introduction to the Theory of Finite Groups.* Edinburgh-London: Oliver & Boyd, 1949. 2nd edition, 1953. 3rd edition, 1957. 4th edition, 1961. Spanish translation, Madrid.

LENTIN, A. and RIVAUD, J. *Eléments d'algèbre moderne,* Paris, 1956. 2nd edition, 1957.

LITTLEWOOD, D. E. *The Theory of Group Characters and Matrix Representations of Groups.* 2nd edition, Oxford: Oxford Univ. Press, 1950.

LJAPIN, E. S. *Semigroups* [in Russian], Moscow, 1960. ✣ English translation, Providence: American Math. Soc., 1964.

LOMONT, J. S. *Applications of Finite Groups.* New York: Academic Press, 1959.

LUGOWSKI, H. and WEINERT, H. J. *Grundzüge der Algebra, Teil* 1, *Allgemeine Gruppentheorie.* Leipzig: Teubner, 1957; *Teil* 2, *Allgemeine Ring- und Körpertheorie.* Leipzig: Teubner, 1958.

MACDUFFEE, C. C. *An Introduction to Abstract Algebra.* New York: Wiley, 1940.

MCCOY, N. H. *Rings and Ideals.* LaSalle (Illinois): Open Court, 1948.

MILLER, G. A., BLICHFELD, H. F., and DICKSON, L. E. *Theory and Applications of Finite Groups.* 2nd edition, New York: Stechert, 1938. Reprint, New York: Dover, 1962.

MILLER, K. S. *Elements of Modern Abstract Algebra.* New York: Harper, 1958.

MOISIL, GR. C. *Introducere in algebra, I, Inele si ideale,* vol. 1. Bucharest, 1954.

MONTEIRO, A. A. *Filtros e ideais* [mimeographed]. Rio de Janeiro: Instituto de Matemática Pura e Aplicada do Conselho Nacional de Pesquinas, 1955.

MONTGOMERY, D. and ZIPPIN, L. *Topological Transformation Groups.* New York: Interscience, 1955.

MORGADO, J. *Elementos de algebra moderna: reticulados, sistemas parcialmente ordenados,* vol. 1. Porto, 1956.

MURNAGHAN, F. D. *The Theory of Group Representations.* Baltimore: Johns Hopkins Press, 1938. Russian translation, Moscow, 1950.

NAKAYAMA, T. and AZUMAYA, G. *Algebra, II, Theory of Rings* [in Japanese], 1954.

NEUMANN, J. VON. *Lectures on Continuous Geometries,* 2 vols. Princeton, 1936-1937. Revised edition: *Continuous Geometry.* Princeton: Princeton Univ. Press, 1960.

NORTHCOTT, D. G. *Ideal Theory.* Cambridge: Cambridge Univ. Press, 1953.

NORTHCOTT, D. G. *Introduction to Homological Algebra.* Cambridge: Cambridge Univ. Press, 1960.

OKUNEV, L. J. *Fundamentals of Modern Algebra [in Russian].* Moscow, 1941.

OSIMA, M. *Theory of Groups* [in Japanese]. 1954.

PICKERT, G. *Projektive Ebenen.* Berlin-Göttingen-Heidelberg: Springer-Verlag, 1955.

PICKERT, G. *Einführung in die höhere Algebra.* Göttingen, 1951.

PLOTKIN, B. I. "Generalized Solvable and Generalized Nilpotent Groups" [in Russian]. UMN 13:4 (1958), 89-172. ❖ English translation in *Translations, Series II,* vol. 17. Providence: American Math. Soc., 1961.

PONTRJAGIN, L. S. *Topological Groups.* Moscow-Leningrad, 1938. 2nd edition, 1954. ❖ English translation, Princeton: Princeton Univ. Press, 1939. Rumanian translation, Bucharest, 1956. ❖ German translation: *Topologische Gruppen,* 2 vols. Leipzig: Teubner, 1957-1958.

POSTNIKOV, M. M. "Definite Families of Functions and Algebras Without Zero Divisors Over the Field of Real Numbers" [in Russian]. UMN 9:2 (1954), 67-104.

POSTNIKOV, M. M. *Fundamentals of Galois Theory* [in Russian]. Moscow, 1960. ❖ English translation: *Fundamentals of Galois Theory.* Groningen: Noordhoff, 1962.

QUEYSANNE, M. and DELACHET, A. *L'algèbre moderne.* Paris: Presses Universitaires de France, 1955. 2nd edition, 1957. Portuguese translation, São Paulo, 1956.

RÉDEI, L. *Algebra,* vol. 1 [in Hungarian]. Budapest, 1954. German translation: *Algebra, Erster Teil.* Leipzig: Geest & Portig, 1959.

REIDEMEISTER, K. *Einführung in die kombinatorische Topologie.* Braunschweig, 1932. Reprint, New York: Chelsea, 1950.

RITT, J. F. *Differential Equations from the Algebraic Standpoint.* New York: American Math. Soc., 1932. 2nd edition, 1947.

RITT, J. F. *Differential Algebra.* New York: American Math. Soc., 1950.

SAMUEL, P. *Algèbre locale* (Mémorial des Sciences Mathématiques, no. 123). Paris: Gauthier-Villars, 1953.

SATO, S. *Theory of Groups* [in Japanese]. Chinese translation, Shanghai, 1934.

SCHILLING, O. F. G. *The Theory of Valuations.* New York: American Math. Soc., 1950.

INDEX

ABELIAN GROUP, 31
  with operator ring, 193
  primary, 204
Abelian Ω-group, 126
Absolute value, 273
Absolutely degenerate, 137
Admissible, Σ-, 195
Admissible set, 199
Albert's theorem, 284
Algebra(s), 221
  automorphism, 108
  automorphism group, 109
  Boolean, 189
  Cayley, 226
  Cayley-Dickson, 242
  complete direct sum, 183
  differential, 8
  generators, 95
  dimension, 224
  division, 221
    real, 225
  [*linear*], factor, 221
  [*universal*], factor, 97
  finite-dimensional, 224
  free, 135
  group, 224
  groupoid, 224
  homological, 8
  [*linear*], homomorphism, 221
  [*universal*], homomorphism, 95
  [*linear*], isomorphism, 221
  [*universal*], isomorphism, 95
  linear, 221
  normed, 283f.
  primitive class, 132
  projective, 8
  quaternion, 225
  rank, 224
  of same type, 95
  semigroup, 224
  subdirect sum, 183
  system of generators, 95
  with unique division, 221
  universal, 94
  of words, 135
Algebraic extension, 315
Algebraic operation. *See* Operation
Algebraic system, 134
All element, 160
α-set, 230
α-subring, 231
Alternating group, 67
Alternative ring, 229
Annihilator, 70, 181, 214
Annihilator isomorphism, 181
Anticommutative law, 36
Anti-isomorphism, 194
Antisymmetry, 17
Archimedean class, 268
Archimedean group, 268
Archimedean ring, 272
Archimedean valuation, 275
Artin's theorem, 229
Ascending chain, 26
Ascending normal series, 122
Associate, 71
Associate decompositions, 75
Associative-commutative ring, 35
  free, 142
Associative law, 30
Associative ring, 35
  free, 142
Associator, 229
Automorphism, 108
  identity, 108
  inner, 109, 110
  Σ-operator, 192
Automorphism group, 109
Axiom of choice, 15

BASIS, 198
  canonical, Cayley algebra, 241
    quaternion algebra, 238
  vector space, 207
Binary algebraic operation, narrow sense, 30
  wide sense (= ternary operation), 30
Binary relation(s), 15
Birkhoff-Witt theorem, 243
Boolean algebra, 189
Boolean lattice, 189
  of sets, 189
Boolean ring, 189
Bound, 26
Bounded on right, 307

CANCELLATION, 31
Cardinal number, 15
Cayley algebra, 226
  canonical basis, 241

Cayley-Dickson algebra, 242
Center, 105, 110
Central isomorphism, 181
Central series, 127
Chain (= linearly ordered set), 21
  ascending, 26
  derived, 129
  lower central, 127
  maximal, 26
  strictly descending, 23
Characteristic of field, 106
Choice, axiom of, 15
Class, archimedean, 268
  primitive, 132
  of partition, 18
Closed element, 289
Closed subset, 290
Closure, 289, 290
Commutative field, 41
Commutative group, 31
Commutative law, 31
Commutator, 123, 302
Commutator group, 123
Compact, 297
  locally, 297
Comparable elements, 21
Complement, of binary relation, 16
  of subset, 13
  of subspace, 205
Complete direct sum, 183
Complete field, 283
Complete inverse image, 107
Complete lattice, 159
Complete system of neighborhoods, 292
  equivalence, 294
Completely regular space, 297
Completion of field with valuation, 283
Component, 173, 182
Composition series, 122
Condition, descending chain, 23
  inductiveness, 23
  maximal, 26
  minimal, 23
Congruence, 96
Conjugate, 315
  Cayley algebra, 227
  quaternion, 226
Conjugate elements, 65
Conjugate subgroups, 65
Connected space, 298
Connection, Galois, 313
Contained in, 21
Contains, 21
Continuity, of inverse, 299
  of multiplication, 299
Continuous homomorphism, 302
Continuum, cardinal number, 15
Convex subset, 266
Correspondence, 14
  semi-linear, 209
Coset(s), 63, 64
  space of, 303
Countable, 15
Cyclic group, 46
  primary, 203
Cyclic subgroup, 43
Cyclic sub-semigroup, 43

DECOMPOSITION(S), 64, 146
  associate, 75
  free, 146
    isomorphism, 153
    refinement, 153
  of Ω-group, 101
Dedekind lattice (= modular lattice), 163, 164, 167
Dedekind ring, 86
Defining relations, 143
Degenerate, absolutely, 137
Degree, extension, 315
  lower, 41
  polynomial, 39, 315
  sum of words, 244
Dense ring, 215
Dependence, linear, 206
Derivation, 249
  inner, 250
Derived chain, 129
Derived group, 125
Derived operation, 134
  principal, 134
Descending chain condition, 23
Dickson. *See* Cayley-Dickson algebra
Differential algebra, 8
Differential ideal, 252
Differential ring, 252, 254
Dimension, (linear) algebra, 224
  vector space, 208
Direct product, 169ff.
Direct sum, 169ff., 178
  complete, 183
  Ω-groups, 186
Direct union, 170
Directed group, 258
Directed set, 258
Directly similar, 174
Disconnected, 298
  locally, 298
Discrete topology, 295
Distinguished subset, 27
Distributive lattice, 163, 187
Distributive laws, 13
Division algebra, 221
Division ring, 59
Divisor, 71
  greatest common, 72
  proper, 75
  of unity, 71
  of zero, 38
Double module, 146

EMBEDDING, 21, 48, 158, 290
Empty relation, 17
Empty set, 13

Endomorphism(s), 110
  Lie ring, 249
  Σ-operator, 192
Endomorphism ring, 113
Equipotent sets, 15
Equivalence of complete systems of neighborhoods, 294
  relative to A, 135
  with respect to double module, 146
Equivalence relation(s), 18
  lattice, 159
  union, 20
Equivalent systems of operations, 134
Euclidean ring, 79
Extension, algebraic, 315
  finite, 315
  normal, 316
Extension field, 315

FACTORS, of normal series, 120
Factor algebra, 97
Factor group, 97
  Ω-, 103
  topology, 303
Factor groupoid, 97
Factor ring, 89, 97, 103
Factor set, 20
Field, 41, 59, 321
  characteristic of, 106
  commutative, 41
  complete, 283
  of complex numbers, as topological field, 307
  extension, 315
  of fractions, 56
  multiplicative group, 41
  p-adic numbers, 117
  prime, 105
  quaternion, as topological field, 307
  of rational fractions, 58
  of rational numbers, p-adic valuation, 282
  of real numbers, as topological field, 307
  residue, 280
  topological, 304
  as universal algebra, 94
  with valuation, completion, 283
    isomorphism, 283
Finite-dimensional (linear) algebra, 224
Finite-dimensional space, 208
Finite extension, 315
Finite order, 44
Finitely generated, 46
Fixed element, 161
Fractions, 52, 55, 56
Fractional ideal, 86
  invertible, 87
Free algebra, 135
Free associative-commutative ring, 142
Free associative ring, 142
Free decompositions, 146
  isomorphism, 153
  refinement, 153
Free element, 131
Free generators, system, 135
Free groupoid, 138
Free R-module, 198
Free product, 144
Free ring, 141
Free semigroup, 139
Free union, 143
Frobenius' theorem, 234
  generalized, 234
Fully invariant subgroups, 194
Fundamental sequence, 283

GALOIS CONNECTION, 313
Galois group, 314
Galois theory, fundamental theorem, 318
Gaussian ring, 77
Gaussian semigroup, 75
Gauss's lemma, 81
Generalized Frobenius' theorem, 234
Generators, 46, 95
  free, 135
Greater than or equal to, 21
Greatest common divisor, 72
Group(s), abelian, 31
    with operator ring, 193
    primary, 204
  additive, of integers, 34
    of rational numbers, 34
  alternating, 67
  archimedean, 268
  automorphism, 109
  commutative, 31
  commutator, 123
  cyclic, 46
  derived, 125
  direct product, 169ff.
  direct sum, 169ff.
  directed, 258
  factor, 97
    topology, 303
  finite, 34
  finitely generated, 46
  Galois, 314
  identical relations for, 131
  with multi-operators, 99
  non-commutative, 34
  Ω-factor, 103
  periodic, 44
  periodic part, 204
  powers of element, 43
  primary, 203
  simple, 67
  symmetric, 34
  topological, 299
  topologization, 303
  torsion-free, 44
  of type p^∞, 114

Group (cont'd)
as universal algebra, 94
See also Ω-groups
Group algebra, 224
Groupoid, 30
factor, 97
free, 138
isomorphic embedding, 48
isomorphism, 46
isotopy, 59
linearly ordered, 255
multiplicative, of ring, 35
with operator semigroup, 192
partially ordered, 255
isomorphism, 256
Σ-operator, 191
topological, 299
unit element, 41
as universal algebra, 138
value, 275
Groupoid algebra, 224
Groupoid ring, integral, 50

HAUSDORFF SPACE, 296
Hausdorff's theorem, 26
Hölder's theorem, 268
Homeomorphism, 294
Homological algebra, 8
Homomorphic image, 95
Homomorphic mapping. *See* Homomorphism
Homomorphism(s), of (linear) algebra, 221
of (universal) algebra, 95
continuous, 302
inversion, 271
monotonic, 266
natural, 97
product of, 95
Σ-operator, 192, 195
sum of, 111
summable, 111
zero, 112
Homomorphism theorem(s), 20, 97

IDEAL(S), 69, 161
(linear) algebra, 221
differential, 252
fractional, 86
invertible, 87
principal, 86
integral, 86
lattice, 161
left, 70
maximal, 85
minimal right, independence of system, 213
Ω-group, 100
one-sided, 70
prime, 85
principal, 78, 86, 160, 199
product, 84
proper, 85
right, 70
semigroup of, 84
Σ-admissible, 195
sum, 101
two-sided, 70
valuation, 280
zero, 69
Idempotent, 189
Identical relations, for groups, 131
in universal algebra, 132
Identity, Jacobi, 36
Identity automorphism, 108
Identity mapping, 14
Identity permutation, 34
Identity relation, 17
Image, homomorphic, 95
inverse, 107
Inclusion, binary relations, 15
sets, 13
Indecomposable, subdirectly, 183
Indecomposable element, 174
Independence, linear, 206
of minimal right ideals, 213f.
Index, of element, 144
of group, 64
Induction, constructions by, 24
Inductiveness, condition, 23
Inner automorphism, 109, 110
Inner derivation, 250
Integers, additive group, 34
p-adic,
ring, 36
Integral domain, 39
Integral group ring, 50
Integral groupoid ring, 50
Integral ideal (= ideal), 86
Integral semigroup ring, 50
Integrally-closed ring, 92
Intersection, binary relations, 15f.
elements of complete lattice, 160
elements of lattice, 155
partitions, 19
sets, 13
Invariant series, 122
Invariant subgroup, 64
Inverse, binary relations, 17
continuity, 299
Inverse element, 32
Inverse image, 107
Inverse isomorphism, 21
Inverse mapping, 15
Inverse permutation, 34
Inversion, 244
Inversion homomorphism, 271
Invertible fractional ideal, 87
Irreducible element, 72
Irreducible representation, 144
Irreducible ring of endomorphisms, 217
Isomorphism, 46
(linear) algebra, 221
annihilator, 181

Isomorphism (cont'd)
central, 181
field with valuation, 283
free decompositions, 153
groupoids, 46
induced by semi-linear correspondence, 210
inverse, partially ordered sets, 21
lattices, 158
normal series, 121
ordered rings, 261
partially ordered groupoids, 256
partially ordered sets, 21
partially ordered sets with closures, 290
sets with closures, 294
Σ-operator, 192, 195
symbol for, 46
topological groups, 302
universal algebras, 95
vector space, 205
Isomorphism theorem, 118
Isotope, principal, 61
Isotopy, 59, 62

JACOBI IDENTITY, 36
Jacobson's theorem, 217
Jordan-Hölder theorem, 122
Jordan ring, 37
Jumps, 268

KERNEL, of homomorphism, 103
Kuratowski-Zorn theorem, 26

LAGRANGE'S THEOREM, 64
Lattice(s), 155
all element, 160
Boolean, 189
of sets, 189
complete, 159
Dedekind (=modular), 163, 164, 167
distributive, 163, 187
of equivalence relations, 159
isomorphic embedding, 158
isomorphism, 158
modular, 163, 164, 167
null element, 160
principal series, 167
of sets, 187
of subalgebras, 156
of subsets, 156
unit element, 160
as universal algebra, 157
Laurent series, 58
Law(s), anticommutative, 36
associative, 30
commutative, 31
distributive, 13
Left annihilator, 70, 214
Left cancellation, 31
Left coset, 63
Left decomposition, 64
Left ideal, 70
Left multiplication, 194
Lemma, Gauss's, 81
Schur's, 217
Zassenhaus', 119
Length, element, 144
element of Gaussian semigroup, 75
normal series, 120
of lattice, 166
product, 232
word, 139
Less than, 21
Levi's theorem, 260
Lie product, 249
Lie ring, 36
of endomorphisms, 249
Limit element, 25
Limit of sequence, 283
Line, 295
rational, 298
Linear, 303
totally, 303
Linear algebra, 221. *See also* Algebra
Linear dependence, 206
Linear space (= vector space), 205
Linear transformation, 209
finite rank, 215
Linearly ordered groupoid, 255
Linearly ordered ring, 261
Linearly ordered set, 21
Locally compact space, 297
Logarithmic valuation, 277
Loop, 59
Lower bound, 26
Lower central chain, 127
Lower degree, 41

MAPPING(S), 14
homomorphic. *See* Homomorphism
identity, 14
inverse, 15
isomorphic. *See* Isomorphism
monotonic, 161
multiplication, 14
natural, 20
one-to-one, 14
Matrix(ices), null, 38
rings of, 38
scalar, 48
square, ring of, 36
Matrix ring, full, 38
Maximal chain, 26
Maximal condition, 26
Maximal element, 26
Maximal ideal, 85
Minimal condition, 23
Minimal element, 22
Models, 18
Modular lattice, 163, 164, 167
Module, double, 146
R-, 193
free, 198
Monotonic homomorphism, 266

Monotonic mapping, 161
Monotonic transformation, 256
Multi-operators, groups with, 99
Multiplication, 30
  continuity, 299
  of fractions, 52
  of homomorphisms, 95
  of mappings, 14
    partial, 14
  of words, 138, 139
  zero, 35
Multiplication table, 223
Multiplicative group, of field, 41
  n-th roots of unity, 34
Multiplicative groupoid, 35
Multiplicative semigroup, 35

n-ARY RELATION, 18
Natural homomorphism, 97
Natural mapping, 20
Natural numbers, additive semigroup, 34
Negative element, of partially ordered set, 256
Neutral, 307
Neighborhoods, 292
  complete system of, 292
Nielsen-Schreier theorem, 153
Nilpotent, 309
  topologically, 307
Nilpotent Ω-group, 127
Non-archimedean valuation, 275
Non-commutative group, 34
Non-commutative semigroup, 34
Norm, in Cayley algebra, 228
  of quaternion, 226
Normal element, 232
Normal extension, 316
Normal form, of words, 245
Normal series, 120
  ascending, 122
  factors, 120
  isomorphism, 121
  of lattice, 166
  length, 120
  refinement, 121
Normal space, 297
Normal subgroup, 64
Normal words, 245
Normed algebra, 283f.
Null, of lattice, 160
Null matrix, 38
Null sequence, 283
Nullary operation, 93

Ω-FACTOR-GROUP, 99, 103
Ω-group (= group with multi-operators Ω), 99
  abelian, 126
  decomposition, 101
  ideal, 100
  nilpotent, 127
  simple, 101
  solvable, 129
Ω-subgroup, 99
Open set, 291
Operation (= algebraic operation), 30
  binary, 93
  derived, 134
  n-ary, 93
  nullary, 93
  partial, 93
  ternary, 93
  unary, 93
Operator domain, 191
Operator endomorphisms, ring of, 197
Operator ring, 195
Opposite element, 33
Order, finite, 44
  of finite group, 34
  infinite, 44

p-ADIC INTEGER, 116
p-adic number, 117
p-adic topologization, 307
p-adic valuation, 282
Partial multiplication of mappings, 14
Partial order, relations of, 20
Partially ordered groupoid, 255
Partially ordered ring, 261
Partially ordered set(s), 21
  with closures, isomorphism, 290
  isomorphism, 21
Partition(s), 18
Periodic group, 44
Periodic part of group, 204
Permutation, 34
Polynomial(s), canonical expression, 80, 83
  degree, 39
  primitive, 80
  ring, 39
Positive elements, of partially ordered group, 256
Power (= cardinal number), 15
Power, of element, 43
Power-associative ring, 229
Power series, ring, 40
Primary abelian group, 204
Primary cyclic group, 203
Prime, 72
  relatively, 74
Prime field, 105
Prime ideal, 85
Primitive class, 132
Primitive polynomial, 80
Principal derived operations, 134
Principal fractional ideal, 86
Principal ideal, 78
  of lattice, 160
Principal ideal ring, 78
Principal isotope, 61
Principal series, 123
  of lattice, 167

Product, of binary relations, 16
direct, 169ff.
ring elements, 262
normal words, 245
Proper divisor, 75
Product, free, 144
ideals, 84
Lie, 249
Projective algebra, 8
Proper ideals, 85
Pseudo-valuation, 278

QUASIFIELD, 59
Quaternion, 225
conjugate, 226
norm, 226
Quaternion algebra, 225
canonical basis, 238

RANK, free abelian group, 200
finite linear transformation, 215
(linear) algebra, 224
Rational fractions, field, 58
Rational line, 298
Rational numbers, additive group, 34
Real division algebra, 224
Recurrence relation, 24
Refinement, direct decomposition, 171
free decomposition, 153
normal series, 121
of lattice, 166
partial ordering, 258
Reflexivity, 17
Regular space, 296
Regular subset, 262
Relation, binary, 15
induced in subset, 18
empty, 17
equivalence, 18
identity, 17
n-ary, 18
of partial order, 20
recurrence, 24
ternary, 18
Relatively prime, 74
Representation, irreducible, 144
Residue-class ring modulo n, 104
Residue field, 280
Right annihilator, 70, 214
Right coset, 64
Right decomposition, 64
Right-hand cancellation, 31
Right ideal, 70
Right multiplication, 194
Ring(s), 35
alternative, 229
anti-isomorphism, 194
archimedean, 272
associative, 35
free, 142
associative-commutative, 35
free, 142
Boolean, 189
Dedekind, 86
dense, of linear transformations, 215
differential, 252
of polynomials, 254
direct sum, 169ff.
division, 59
endomorphism, 113
euclidean, 79
factor, 89, 97, 103
finitely generated, 46
of fractions, 56
free, 141
of functions, full, 39
Gaussian, 77
of integers, 35
integral, group, 50
integral groupoid, 50
integral semigroup, 50
integrally-closed, 92
irreducible, of endomorphisms, 217
isomorphic embedding, 48
isomorphism, 46
isotopy, 62
Jordan, 37
Lie, 36
of endomorphisms, 249
linearly ordered, 261
of matrices, 38
multiplicative groupoid, 35
multiplicative semigroup, 35
of operator endomorphisms, 197
operator, 195
ordered, 261
of p-adic integers, 116
partially ordered, 261
of polynomials, 39
of power series, 40
power-associative, 229
product of elements, 262
principal ideal, 78
residue-class, modulo n, 104
Σ-operator, 195
simple, 69
of square matrices, 36
topological, 304
unit element, 41
as universal algebra, 141
with valuation, 273
valuation, 274, 278
of vectors, 36
zero, 35
R-module, 193
Roots of unity, multiplicative group, 34

SCALAR MATRICES, 48
Schreier's theorem, 121
Schur's lemma, 217
Semigroup, 30
of endomorphisms, 110
finitely generated, 46
free, 139

Semigroup (cont'd)
Gaussian, 75
of ideals, 84
multiplicative, of ring, 35
of negative elements, 257
non-commutative, 34
of non-zero fractional ideals, 87
operator, 192
of positive elements, 257
symmetric, 34
topological, 299
as universal algebra, 139
Semigroup algebra, 224
Semigroup ring, integral, 50
Semi-linear correspondence, 209
Sequence, fundamental, 283
limit of, 283
null, 283
Series, central, 127
composition, 122
invariant, 122
normal, 120
principal, 123, 167
solvable, 129
Sets, admissible, 199
Boolean lattice, 189
with closures, isomorphism of, 294
directed, 258
empty, 13
equipotent, 15
factor, 20
lattice, 187
open, 291
ordered, 21
linearly, 21
partially ordered, 21
See also Ordered set, Partially ordered set, Well-ordered set
Σ-admissible, 195
Σ-admissible subgroupoid, 191
Σ-operator automorphism, 192
Σ-operator groupoid, 191
Σ-operator homomorphism, 192, 195
Σ-operator isomorphism, 192, 195
Σ-operator ring, 195
Similar, directly, 174
Simple group, 67
Simple Ω-group, 101
Simple ring, 69
Single-valued function, 14
Skewfield, 321
Solvable series, 129
Solvable Ω-group, 129
Space, compact, 297
locally, 297
completely regular, 297
connected, 298
of cosets, 303
finite-dimensional, 208
Hausdorff, 296
linear (= vector space), 205
locally compact, 297
normal, 297
regular, 296
topological, 290f.
totally disconnected, 298
vector, 205
Square, of ring, 125
of set, 15
Subalgebra, 94
Subalgebra (linear), 221
Subdirect sum, 183
Subdirectly indecomposable, 183
Subfield, 43
generated by subset, 45
Subgroup, 43, 194
characteristic, 194
conjugate, 65
cyclic, 43
fully invariant, 194
generated by subset, 45
generated by subgroups, 46
invariant, 64
fully, 194
normal, 64
Ω-, 99
unit, 43
Subgroup theorem, 146
Subgroupoid, 42, 94
Σ-admissible, 191
Sublattice, 158
Submodule, 193
Subring, 43,
differential, 252
generated by element, 45
generated by subset, 45
generated by subrings, 46
Σ-admissible, 195
zero, 43
Sub-semigroup, 42
cyclic, 43
generated by subset, 45
Subset(s), 13
closed, 290
closure, 290
complement, 13
convex, 266
distinguished, 27
lattice, 156
regular, 262
Subspace, complement, 205
linear, 205
topological, 295
Subword, 132
Sum, complete direct, 183
direct, 169ff., 178, 186
of fractions, 55
of homomorphisms, 111
of ideals, 101
of infinite set of homomorphisms, 111
of products in ring, 262
subdirect, 183
of words, 141, 244
Symmetric group, 34

Symmetric semigroup, 34
Symmetry, 17
System, algebraic, 134
  of defining relations, 143
  equivalent, of operations, 134
  of generators, 46, 95, 135
  of neighborhoods, 292

TERNARY OPERATION, 93
Ternary relation, 18
Theorem, Albert's, 284
  Artin's, 229
  Birkhoff-Witt, 243
  Frobenius', 234
    generalized, 234
  fundamental, of Galois theory, 318
  Hausdorff's, 26
  Hölder's, 268
  homomorphism, 97
  isomorphism, 118
  Jacobson's, 217
  Jordan-Hölder, 122
  Kuratowski-Zorn, 26
  Lagrange's, 64
  Levi's, 260
  Nielsen-Schreier, 153
  of primitive element, 315
  Schreier's, 121
  subgroup, 146
  Zermelo's, 26
Theorems, homomorphism, 20
Topological field, 304
Topological groups, 299
  isomorphism, 302
Topological groupoid, 299
Topological ring, 304
Topological semigroup, 299
Topological space, 290f.
Topologically nilpotent, 307
Topologization, 299
  field of complex numbers, 307
  linear, 303
  p-adic, 307
  quaternion field, 307
  real number field, 307
  totally linear, 303
Topology, 290
  discrete, 295
Torsion-free group, 44
Totally disconnected space, 298
Transform, 109
Transformation, 34
  linear, 209
    of finite rank, 215
  monotonic, 256
Transitivity, 17
Translation, 300
Trivial closure, 289
Trivial partial order, 21
Trivial valuation, 275
Type p^∞, 114

UNARY OPERATION, 93
Union, of binary relations, 15f.
  direct, 170
  elements of complete lattice, 160
  elements of lattice, 155
  equivalence relations, 20
  free, 143
  sets, 13
Unit element, group, 32
  groupoid, 41
  lattice, 160
  in multiplication of mappings, 14
  ring, 41
Unit subgroup, 43
Unitary R-module, 193
Unity, divisor of, 71
Universal algebra 94. *See also* Algebra
Upper bound, 26

VALUATION, archimedean, 275
  logarithmic, 277
  non-archimedean, 275
  p-adic, 282
  pseudo-, 278
  ring with, 273
  trivial, 275
Valuation ideal, 280
Valuation ring, 274, 278
Value, 273
  absolute, 273
Value groupoid, 275
Vector space, basis, 207
  dimension, 208
  isomorphism, 205
  over fields, 205
Vectors, ring of, 36

WAERDEN, VAN DER, 7,
Well-ordered set, 25
Witt, theorem of Birkhoff-, 243
Word(s), 131, 138, 145, 244
  addition, 141
  algebra, 135
  empty, 139, 145
  equivalent relative to A, 135
  length, 139
  normal, 245
  sums, 244

ZASSENHAUS' LEMMA, 119
Zermelo's theorem, 26
Zero, 33
  divisors of, 38
Zero function, 39
Zero homomorphism, 112
Zero ideal, 69
Zero multiplication, 35
Zero ring, 35
Zero subring, 43